

STIC-Biotech/ChemLib

140214

From: Chan, Christina
Sent: Monday, December 13, 2004 4:29 PM
To: Basi, Nirmal; STIC-Biotech/ChemLib
Subject: RE: Rush search for 09/995,225

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Basi, Nirmal
Sent: Monday, December 13, 2004 1:49 PM
To: Chan, Christina
Subject: Rush search for 09/995,225

Examiner: Nirmal S. Basi
Art Unit 1646
Office: Remsen Building, Room 4D68
Mail Room: Remsen Building, room 4C70

Sequence search:

App. #: 09/995,225
Result format: Paper.

Title: **ENDOGENOUS AND NON-ENDOGENOUS VERSIONS OF HUMAN G PROTEIN-
COUPLED RECEPTORS**

Inventors: Ruoping Chen et al

Priority Date: 10/13/98

Please search:

- i) Please search SEQ ID NO: 1, 2, 15, 16, 41 and 42,
- ii) Polynucleotide encoding the polypeptide of SEQ ID NO: 16

Search issued and commercial databases.

Thanks,

RECEIVED
DEC 13 2005
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: 12/14/04
Date Completed: 12/15/04
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search 4
NA Sequence: # _____
AA Sequence: # 271
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: plot
WWW/Internet: _____
Other(Specify): _____

revert to
NA

This Page Blank (uspto)

Gencore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:34:50 ; Search time 3576.97 Seconds

(without alignments)
14040.270 Million cell updates/sec

Title: US-09-995-225b-15

Perfect score: 1062

Sequence: 1 atgagagcacagcagccca.....ctataaagatcccccgtga 1062

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_ncg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_str:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1062	100.0	1062	6	BD183204
2	1062	100.0	1062	6	AX498192
3	1062	100.0	1062	6	BD103852
4	1061.6	100.0	1167	9	AY635179
5	1060.4	99.8	1062	6	BD183203
6	1060.4	99.8	1062	6	BD103851
7	1058.8	99.7	1343	6	AX254977
8	1057.4	99.6	1059	6	AX453227
9	1050.8	98.9	1062	6	AX664703
10	1050.8	98.9	1526	6	AX664701
11	1048.4	98.7	1119	6	AX254975
12	1048.4	98.7	1130	6	AX375235
13	945.2	89.0	1062	6	AX463235
14	940.2	88.5	957	6	CQ737667
15	938	88.3	11000	2	AL139235_0
16	936.4	88.2	1002	6	AX375442
17	936.4	88.2	1158	6	BD144286
18	936.4	88.2	1167	9	AB083594
19	936.4	88.2	101882	9	AC021089

20	936	88.1	1032	6	AX463231	AX463231 Sequence
21	936	88.1	1070	6	AX463230	AX463230 Sequence
22	936	88.1	1826	6	AX463232	AX463232 Sequence
23	934.8	88.0	1000	6	AX342465	AX342465 Sequence
24	934.8	88.0	1788	6	CQ739127	CQ739127 Sequence
25	842.6	79.3	864	6	AX230119	AX230119 Sequence
26	809.6	76.2	963	6	AX657444	AX657444 Sequence
27	788.4	74.2	183326	10	AC124438	AC124438 Mus muscu
28	779.6	73.4	140885	2	AC119798	AC119798 Rattus no
29	779.6	73.4	203372	2	AC121346	AC121346 Rattus no
30	779.6	73.4	230813	2	AC114097	AC114097 Rattus no
31	758	71.4	795	6	AX549301	AX549301 Sequence
32	758	71.4	795	6	AX549301	AX549301 Sequence
33	555.4	52.3	558	9	AY255545	AY255545 Homo sapi
34	524	49.3	930	6	AX230116	AX230116 Sequence
35	444.4	41.8	1149	5	AY288414	AY288414 Takifugu
36	268.2	25.3	945	9	AY255622	AY255622 Homo sapi
37	268.2	25.3	1962	6	AX375239	AX375239 Homo sapi
38	266.6	25.1	1122	6	BD185294	BD185294 Novel G P
39	266.6	25.1	1122	6	BD140885	BD140885 Novel G P
40	266.6	25.1	1125	6	AX498182	AX498182 Sequence
41	266.6	25.1	1125	6	AX664706	AX664706 Sequence
42	266.6	25.1	1389	9	AY288421	AY288421 Homo sapi
43	266.6	25.1	1719	6	AX664704	AX664704 Sequence
44	265.6	25.0	951	6	AX657458	AX657458 Sequence
45	245.8	23.1	1113	5	AY288413	AY288413 Takifugu

ALIGNMENTS

RESULT 1

BD183204

LOCUS

DEFINITION Novel G protein-coupled receptor and its DNA.

ACCESSION BD183204

VERSION BD183204.1 GI:31875404

KEYWORDS JP 2002345481-A/4.

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

AUTHORS Terao,Y., Matsui,H. and Shintani,Y.

TITLE Novel G protein-coupled receptor and its DNA

JOURNAL Patent: JP 2002345481-A 4 03-DEC-2002;

COMMENT TAKEDA CHEMICAL INDUSTRIES LTD

OS Homo sapiens (human)

PN JP 2002345481-A/4

PD 03-DEC-2002

PF 01-JUN-2001 JP 2001166688

PI YASUKO TERAO,HIDEKI MATSUI,YASUSHI SHINTANI

PC C12N15/09,A61K45/00,A61P1/00,A61P3/00,A61P9/00,A61P25/00, PC

A61P29/00,

PC A61P35/00,A61P37/00,C07K14/705,C07K16/28,C12N1/15,C12N1/19, PC

C12N1/21,

PC C12N5/10,C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC

G01N33/53,

PC G01N33/566,C12N15/00,C12N5/00

CC Novel G protein-coupled receptor and its DNA

FT

FT source

FEATURES

source

1..1062

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1062; DB 6; Length 1062;

Best local Similarity 100.0%; Pred. No. 5.2e-247;

Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGACACGACGCGCCACCTGCGAGCAAGCTGCTGCTTGTGTGTCCTCCCGGC 60
DB 1 ATGGAGACACGACGCGCCACCTGCGAGCAAGCTGCTGCTTGTGTGTCCTCCCGGC 60
QY 61 TCGGCTGCGGCTTGGGTTTCGTCGCGCGGTGCTACTAGACGCTTGTGTCCTCGGT 120
DB 61 TCGGCTGCGGCTTGGGTTTCGTCGCGCGGTGCTACTAGACGCTTGTGTCCTCGGT 120
QY 121 TTAACGCAAAATATCTTGACAGTATCTCTCCAGCTGGTGCAAGAGACGAAG 180
DB 121 TTAACGCAAAATATCTTGACAGTATCTCTCCAGCTGGTGCAAGAGACGAAG 180
QY 181 TCTCTACAACTATCTTGGACTGCTGCGCGGACATCTTGGTCTCTTTTTCATA 240
DB 181 TCTCTACAACTATCTTGGACTGCTGCGCGGACATCTTGGTCTCTTTTTCATA 240
QY 241 GTGTTGTGGAATCTCTGTGGAGATTTCATCTTGAACATGAGATGCTCAGTCCGC 300
DB 241 GTGTTGTGGAATCTCTGTGGAGATTTCATCTTGAACATGAGATGCTCAGTCCGC 300
QY 301 GACAAATATGAAATGCTGGAATCTCATCTCCACACCTCCATATGATTAAGTGA 360
DB 301 GACAAATATGAAATGCTGGAATCTCATCTCCACACCTCCATATGATTAAGTGA 360
QY 361 CCGTTAACCATGACAGATATACGCTGCTGCGACCGGCTCAAGTACACACGCTCA 420
DB 361 CCGTTAACCATGACAGATATACGCTGCTGCGACCGGCTCAAGTACACACGCTCA 420
QY 421 TACCCAGCCGACCCGGAAGTCAATGTAGTGTATACATCAGTCACTGCTCTGACAC 480
DB 421 TACCCAGCCGACCCGGAAGTCAATGTAGTGTATACATCAGTCACTGCTCTGACAC 480
QY 481 ATCCCTATTAATCTGTGCGCCCAACATCTGAGTGAAGATTAATCAGACCTCTGTGCAT 540
DB 481 ATCCCTATTAATCTGTGCGCCCAACATCTGAGTGAAGATTAATCAGACCTCTGTGCAT 540
QY 541 CAGCTCTCATCTGATTCAGTCTCAACGCTGCTGAGCGGCTCCACATCTTCTTC 600
DB 541 CAGCTCTCATCTGATTCAGTCTCAACGCTGCTGAGCGGCTCCACATCTTCTTC 600
QY 601 ATCTTGAATCAATCTGTGTGTAACAAGCTCAAGAGAGAGCAATTTTGTCTCCGTGC 660
DB 601 ATCTTGAATCAATCTGTGTGTAACAAGCTCAAGAGAGAGCAATTTTGTCTCCGTGC 660
QY 661 TACTCAGCGGAGAACACCGCACTTGTTCACCATTAATCTTCATCTTTCACACATT 720
DB 661 TACTCAGCGGAGAACACCGCACTTGTTCACCATTAATCTTCATCTTTCACACATT 720
QY 721 TGGGCCCCCGCATCATGATCTTTACACCTCTATGAGGCGCCCATCCAGAACCGC 780
DB 721 TGGGCCCCCGCATCATGATCTTTACACCTCTATGAGGCGCCCATCCAGAACCGC 780
QY 781 TGGCTGTGACATCATGTCGACATGTGCAATGTGCAATGTGCAATGTGCAATGTG 840
DB 781 TGGCTGTGACATCATGTCGACATGTGCAATGTGCAATGTGCAATGTGCAATGTG 840
QY 841 AACTCTTCTCTACTGCTTCAATCAGCAAGCGTTCGCAACATGTGCAAGCGCCACGCTC 900
DB 841 AACTCTTCTCTACTGCTTCAATCAGCAAGCGTTCGCAACATGTGCAAGCGCCACGCTC 900
QY 901 AAGGCTTCTCTCAATGTCGAGAACCTGTACAGTCTTACACCAATCATTAATCTTCC 960
DB 901 AAGGCTTCTCTCAATGTCGAGAACCTGTACAGTCTTACACCAATCATTAATCTTCC 960
QY 961 ATAAAGATGAGCCCTGATCTCGCGGAACTCACTGATCAAGATGTGTGTATAC 1020
DB 961 ATAAAGATGAGCCCTGATCTCGCGGAACTCACTGATCAAGATGTGTGTATAC 1020
QY 1021 CAGTATGACAAAATGAGAAAACCTATAAAGTATCCCGTGA 1062
DB 1021 CAGTATGACAAAATGAGAAAACCTATAAAGTATCCCGTGA 1062

RESULT 2
AX498192 1062 bp DNA linear PAT 26-SHP-2002
LOCUS
DEFINITION
Sequence 15 from Patent WO0242461.
ACCESSION
AX498192
VERSION
AX498192.1 GI:23343117
KEYWORDS
synthetic construct
synthetic construct
artificial sequences.
SOURCE
ORGANISM
REFERENCES
1
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1062
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Novel Sequence"
ORIGIN
Query Match 100.0%; Score 1062; DB 6; Length 1062;
Best Local Similarity 100.0%; Pred. No. 5.2e-247;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGACACGACGCGCCACCTGCGAGCAAGCTGCTGCTTGTGTGTCCTCCCGGC 60
DB 1 ATGGAGACACGACGCGCCACCTGCGAGCAAGCTGCTGCTTGTGTGTCCTCCCGGC 60
QY 61 TCGGCTGCGGCTTGGGTTTCGTCGCGCGGTGCTACTAGACGCTTGTGTCCTCGGT 120
DB 61 TCGGCTGCGGCTTGGGTTTCGTCGCGCGGTGCTACTAGACGCTTGTGTCCTCGGT 120
QY 121 TTAACGCAAAATATCTTGACAGTATCTCTCCAGCTGGTGCAAGAGACGAAG 180
DB 121 TTAACGCAAAATATCTTGACAGTATCTCTCCAGCTGGTGCAAGAGACGAAG 180
QY 181 TCTCTACAACTATCTTGGACTGCTGCGCGGACATCTTGGTCTCTTTTTCATA 240
DB 181 TCTCTACAACTATCTTGGACTGCTGCGCGGACATCTTGGTCTCTTTTTCATA 240
QY 241 GTGTTGTGGAATCTCTGTGGAGATTTCATCTTGAACATGAGATGCTCAGTCCGC 300
DB 241 GTGTTGTGGAATCTCTGTGGAGATTTCATCTTGAACATGAGATGCTCAGTCCGC 300
QY 301 GACAAATATGAAATGCTGGAATCTCATCTCCACACCTCCATATGATTAAGTGA 360
DB 301 GACAAATATGAAATGCTGGAATCTCATCTCCACACCTCCATATGATTAAGTGA 360
QY 361 CCGTTAACCATGACAGATATACGCTGCTGCGACCGGCTCAAGTACACACGCTCA 420
DB 361 CCGTTAACCATGACAGATATACGCTGCTGCGACCGGCTCAAGTACACACGCTCA 420
QY 421 TACCCAGCCGACCCGGAAGTCAATGTAGTGTATACATCAGTCACTGCTCTGACAC 480
DB 421 TACCCAGCCGACCCGGAAGTCAATGTAGTGTATACATCAGTCACTGCTCTGACAC 480
QY 481 ATCCCTATTAATCTGTGCGCCCAACATCTGAGTGAAGATTAATCAGACCTCTGTGCAT 540
DB 481 ATCCCTATTAATCTGTGCGCCCAACATCTGAGTGAAGATTAATCAGACCTCTGTGCAT 540
QY 541 CAGCTCTCATCTGATTCAGTCTCAACGCTGCTGAGCGGCTCCACATCTTCTTC 600
DB 541 CAGCTCTCATCTGATTCAGTCTCAACGCTGCTGAGCGGCTCCACATCTTCTTC 600
QY 601 ATCTTGAATCAATCTGTGTGTAACAAGCTCAAGAGAGAGCAATTTTGTCTCCGTGC 660
DB 601 ATCTTGAATCAATCTGTGTGTAACAAGCTCAAGAGAGAGCAATTTTGTCTCCGTGC 660

OY	661	TACTCAGGGGGAACAACACGGCCATCTGTGTTACATTAATCTGCATCTTTGGACAATT	720
Db	661	TACTCAGGGGGAACAACACGGCCATCTGTGTTACATTAATCTGCATCTTTGGACAATT	720
OY	721	TGGGCCCCCGCATGATCATGATTTCTTTACGACCTGTATGGGGCGCCCATCGAAGCCG	780
Db	721	TGGGCCCCCGCATGATCATGATTTCTTTACGACCTGTATGGGGCGCCCATCGAAGCCG	780
OY	781	TGGCTGTGTCACATCATGTCCGACATTGGCCAAATGCTTAGCCCTTTTGAAACAAGCCATC	840
Db	781	TGGCTGTGTCACATCATGTCCGACATTGGCCAAATGCTTAGCCCTTTTGAAACAAGCCATC	840
OY	841	AACTCTTCCTACTGCTTCATCAGCAAGGGGTTCCGACATGAGCGCCGACAGCTC	900
Db	841	AACTCTTCCTACTGCTTCATCAGCAAGGGGTTCCGACATGAGCGCCGACAGCTC	900
OY	901	AAGCTTCTTCAAGTGCAGAGCAACCTGTACAGTCTTACCAACATCATTAATCTTCC	960
Db	901	AAGCTTCTTCAAGTGCAGAGCAACCTGTACAGTCTTACCAACATCATTAATCTTCC	960
OY	961	ATTAACAGTAGCCCTCTGATCTCCGCGGCAAACTTCACTGCATTCAGATCTGTGTAC	1020
Db	961	ATTAACAGTAGCCCTCTGATCTCCGCGGCAAACTTCACTGCATTCAGATCTGTGTAC	1020
OY	1021	CAGTATGCAAAAATGSAAAACCTATAAAGTATCCCGCTGA	1062
Db	1021	CAGTATGCAAAAATGSAAAACCTATAAAGTATCCCGCTGA	1062

RESULT 3	BD103852	LOCUS	BD103852	1062 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Novel G protein-coupled receptor and its DNA.						
ACCSSION	AF103852						

ACCESSION	BDJ03852	GI:22649426
VERSION	BDJ03852.1	
KEYWORDS	MO_0194582-A/4.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCES
1 (bases 1 to 1062)
Terao, Y., Matsui, H. and Shintani, Y.
Novel G protein-coupled receptor and its DNA
Patent: WO 0194582-A 4 13-DEC-2001;

COMMENT	OS	Homo sapiens (human)
	OS	Homo sapiens (human)

PD 13-DEC-2001 WO 2001JP004643
 PD 01-JUN-2001 WO 170446.23 -JUN-2000 JP 00P 194926 PI
 PR 02-JUN-2000 JP 00P 170446.23 -JUN-2000 JP 00P 194926 PI
 YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
 PC C12N5/12, C12P21/02, C07K14/705, C07K16/28, A61K45/00, A61P25/00,
 A61P29/00,
 PC A61P37/00, A61P35/00, A61P33/00, A61P31/00, G01N33/566, PC
 G01N33/50,

```
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT 1..1062
FT /organism='Homo sapiens (human)'.
FEATURES
    source 1..1062
```

ORIGIN

Query Match	100.0%	Score 1062;	DB 6;	Length 1062;
Best Local Similarity	100.0%	Pred. No. 5, 2e-247;		
Matches 1062;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 ATGGAGCACACGCGCCCACTCGCAGCCACAGCTGGCTTGGTGTCCCCGGC 60

Db

1 ATGAGCACAAGCAGCCCACTCGACGACAAGCTCGCTCTTGGTGTCCCGGAGC 60

Qy 61 TCGGCTGCGGCTTTGGTTTCGTCCCGGTGGTCTACTACAGCCTCTTGGTGCCTGGT 120
Db 61 TCGGCTGCGGCTTTGGTTTCGTCCCGGTGGTCTACTACAGCCTCTTGGTGCCTGGT 120

Qy 121 TTACAGAAATATTCTGACGTATCATCTCCAGCTGGTGGCAAGAACAAG 180
 |||||
Db 121 TTACCGAATATCTTGACGTATCATCTCCAGCTGGTGGCAAGAACAAG 180
 |||||

Accession	Sequence	Length
07	TCCTCTACCAACATCTCTTGGCACTGGCTGCGCCGCAACCTTGGTCTCTTTTATATA	240
Db	TCCTCTACCAACATCTCTTGGCACTGGCTGCGCCGCACTCTTGGTCTCTTTTATATA	240

301 GACAAGATCATAGAAGTGCCTGAATTTCTCATTCATCCATCCACACTTCATATGGAATTTACTGTA 360

Db 301 GACAAATCATAGAAATGCTGGAAATTCATCCATCAACCTTCATATGAGATTACTGTA 360

Dy 361 CCGTTAACCAATGCACAGGTATATGCTGCTGCCACCGCTCAAATGACACACAGGTTCA 420

DB	QY
361	421
CCGTTAACCATTTGACAGGTATATGCTGTCTGCACCCGCTCAAGTACCAACAGGTCTCA	TACCCAGCCGACACCCGAAAGTCATTGTTAAGTGTTCATCATCAGCTCTTCCTGACGAC
420	480

Db 421 TACCAGGCCGACCCGGAAAGTATTGTAGTGTTCATCACCTGCTTCGACGACC 480

Qy 481 ATCCCTATTTCGTGTGGCCGACATTTGAGCTGAGAGCTACATCAGCACTCTGTGCAAT 540

DB 481 ATCCCGATTACCTGGTGGCCGACATCTGGACATGAGACATACATACACATCTGTCGAT 540

QY 541 CAGCTGCTCATCTGGATTCACATCTGTTTCAACGCTTACCTGGTGCCTGCTCCATCTTCTC 600

601 ATCTTGAACCTCAATCATTTGTGTACAGCTGACGAGGAGAACGCAATTTTTCGTCCTCCGCGGC 660
 601 ATCTTGAACCTCAATCATTTGTGTACAGCTGACGAGGAGAACGCAATTTTTCGTCCTCCGCGCGGC 660
 601 ATCTTGAACCTCAATCATTTGTGTACAGCTGACGAGGAGAACGCAATTTTTCGTCCTCCGCGCGGC 660

Qy	Db
661	661
TACTCCAGGGGAAGACACCGGCATCTTGTTCACATTA	TACTCCAGGGGAAGACACCGGCATCTTGTTCACATTA
CTTCATCTTTGCCACACTT	CTTCATCTTTGCCACACTT
720	720
661	661
TACTCCAGGGGAAGACACCGGCATCTTGTTCACATTA	TACTCCAGGGGAAGACACCGGCATCTTGTTCACATTA
CTTCATCTTTGCCACACTT	CTTCATCTTTGCCACACTT
720	720

Qy	Db
721 TGGGCCCCCGCATCATCATGATTCTTTACCACTCTATGGGGGCCCATTCGAACCGC 780	721 TGGGCCCCCGCATCATCATGATTCTTTACCACTCTATGGGGGCCCATTCGAACCGC 780

Qy 781 TGGCTGGTGCATCATGTCCGACATTGCCAATGCTGACCCCTTCTGAAACAGCCATC 840

Db 781 TGGCTGGTGCATCATGTCCGACATTGCCAATGCTGACCCCTTCTGAAACAGCCATC 840

97	AACTTCTTCCTCTACGCTCATCAGCAAGCGGTTCCGACACATGGCAGCGGCACGCTC	900
Db	AACTTCTTCCTCTACGCTCATCAGCAAGCGGTTCCGACACATGGCAGCGGCACGCTC	900

961 **ATGACCTTTCACGCTGACCTTACACCAATCTTATCTTTTCC** 960
 Db 901 **ATGACCTTTCACGCTGACCTTACACCAATCTTATCTTTTCC** 960
 Oy 90 **ATGACCTTTCACGCTGACCTTACACCAATCTTATCTTTTCC** 960
 Oy **ATGACCTTTCACGCTGACCTTACACCAATCTTATCTTTTCC** 960

Db 961 ATAAAGATAGCCCTGGATCTGCGCGGCAAACTCAACATGCATCAAGATCTGGTGTAC 1020
 1021 CAGTATGACAAAAATGSAAAACCTATTAAGATATCCCCGTGA 1062
 Oy

Db 1021 CAGTATGACAAAATGSAALACTATATAAGTATCCCCCGGA 1062

RESULT 4

AY635179 1167 bp mRNA linear PRI 04-JUN-2004
LOCUS AY635179
DEFINITION Homo sapiens G protein-coupled receptor 139 (GPR139) mRNA, complete
ACCESSION AY635179
VERSION AY635179.1 GI:49413551
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Bonner,T.I., Nagle,J.W. and Kauffman,D.
TITLE Complete coding sequence of GPR139
JOURNAL Unpublished
REFERENCE
AUTHORS Bonner,T.I., Nagle,J.W. and Kauffman,D.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2004) Lab of Genetics, NIMH, Bldg 36, Rm 3D06,
MSC4094, Bethesda, MD 20892-4094, USA
FEATURES
source
1. .1167
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p12.3"
/tissue="brain"
1. .1167
/gene="GPR139"
92. .1153
/gene="GPR139"
/product="G protein-coupled receptor 139"
/protein_id="A465818.1"
/db_xref="GI:49413552"
/translation="MEHTYAHLAANSLSWBSGACGLGVVYVYSLILCLGHPAN
ILTVILISOLVARROKSSIVYLAADILVFTVDFLEDFILMOMPOYDK
ILVLESSLHISWITVPLTIDRYLAADILVFTVDFLEDFILMOMPOYDK
IPYWMNPWTEDEYISTSVHVLWICHFTVYLVPCSPFINSIIVGLRKSFRLL
RGYSTGTALITLITSIPTALWAPRIIMLVHVGAPIONWLVHMSDINMALL
NPAINFVLCFIEKFRPMMAATLKAFFGCKQKQPVQFTNHFSTISSEWISPAHSHC
IKMLVVOYDKNGPKIKVSP"
complement (1142..1167)
ORIGIN
primer_bind
Query Match 100.0%; Score 1061.6; DB 9; Length 1167;
Best Local Similarity 99.9%; Pred. No. 6.5e-247;
Matches 1061; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCAGACGACGCCCACTCGACGCAACAGCTGCTGTCTTGTGTGTCGCCCGGC 60
DB 92 ATGGAGCAGACGACGCCCACTCGACGCAACAGCTGCTGTCTTGTGTGTCGCCCGGC 151
QY 61 TCGGCTGCGGCTTGGGTTTCGTGCGCCGCTGTCTACTACAGCTCTGTGCTGCTCGGT 120
DB 152 TCGGCTGCGGCTTGGGTTTCGTGCGCCGCTGTCTACTACAGCTCTGTGCTGCTCGGT 211
QY 121 TTACGACGCAATATCTTGACAGTATCATCTCTCCAGCTGTGGCAAGAGAGAG 180
DB 212 TTACGACGCAATATCTTGACAGTATCATCTCTCCAGCTGTGGCAAGAGAGAG 271
QY 181 TCCCTCTCAACTATCTCTTGACACTCGCTGTGCGCAATCTTGTGCTCTTTTTCATA 240
DB 272 TCCCTCTCAACTATCTCTTGACACTCGCTGTGCGCAATCTTGTGCTCTTTTTCATA 331
QY 241 GTGTTTGGGACTTCTCTTGGAGATTTCATCTTGAACATGACAGATGCCCTCAGTCCC 300
DB 332 GTGTTTGGGACTTCTCTTGGAGATTTCATCTTGAACATGACAGATGCCCTCAGTCCC 391
QY 301 GACAGATCATGAGAGTCTGGAATCTCATCCATCCACAGCTCCATATGATTAATCTGTA 360
DB 392 GACAGATCATGAGAGTCTGGAATCTCATCCATCCACAGCTCCATATGATTAATCTGTA 451

QY 361 CGGTTAACCATTTGACAGGTATATCGCTGTGCGCAACCGCTCAAGTACACAGGTTCA 420
DB 452 CGGTTAACCATTTGACAGGTATATCGCTGTGCGCAACCGCTCAAGTACACAGGTTCA 511
QY 421 TACCCAGCCCGACCCGGAAGTCAATTTAGTGTTTACATCAGCTGCTCTGACAGC 480
DB 512 TACCCAGCCCGACCCGGAAGTCAATTTAGTGTTTACATCAGCTGCTCTGACAGC 571
QY 481 ATCCCTATTACTGTGGGGCCAAACATCTGAGATGTAAGACTATGACACCTGTGCAT 540
DB 572 ATCCCTATTACTGTGGGGCCAAACATCTGAGATGTAAGACTATGACACCTGTGCAT 631
QY 541 CAGTCTCATCTGATCCATCCATGCTTCAACCGTCTACCTGAGTGCCTGCTCATCTTTC 600
DB 632 CAGTCTCATCTGATCCATCCATGCTTCAACCGTCTACCTGAGTGCCTGCTCATCTTTC 691
QY 601 ATCTTGAATCAATCATTTGTGTACAGAGTCAAGAGGAGAACATTTGTCTCGTGGC 660
DB 692 ATCTTGAATCAATCATTTGTGTACAGAGTCAAGAGGAGAACATTTGTCTCGTGGC 751
QY 661 TACTCCAGGGGAGAACACCGCCATCTGTTCACCAATTACTGTCATCTTGGCACATT 720
DB 752 TACTCCAGGGGAGAACACCGCCATCTGTTCACCAATTACTGTCATCTTGGCACATT 811
QY 721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTTATAGGGGCGCCATCCAGAACCG 780
DB 812 TGGGCCCCCGCATCATCATGATTTCTTTACCACTTATAGGGGCGCCATCCAGAACCG 871
QY 781 TGGCTGTGACATCATGTCGACATTTGCCAAGTGTAGCCCTTCTGAAACAGCCATC 840
DB 872 TGGCTGTGACATCATGTCGACATTTGCCAAGTGTAGCCCTTCTGAAACAGCCATC 931
QY 841 AACTTCTTCTCTACTGCTTCATCAGCAAGGTTCCGACCATGAGCGCGCACGCTC 900
DB 932 AACTTCTTCTCTACTGCTTCATCAGCAAGGTTCCGACCATGAGCGCGCACGCTC 991
QY 901 AAGGCTTCTTCAAGTCCGAGAACCACTGTACAGTTCTTACACCAATCAATCAATTTTC 960
DB 992 AAGGCTTCTTCAAGTCCGAGAACCACTGTACAGTTCTTACACCAATCAATCAATTTTC 1051
QY 961 ATTAACAGTACGCCCTGATCTGCGCGGCAACCTCAACATGATCAAGATGCTGTGTAC 1020
DB 1052 ATTAACAGTACGCCCTGATCTGCGCGGCAACCTCAACATGATCAAGATGCTGTGTAC 1111
QY 1021 CAGTATGACAAAATGGAACCTATTAAGATATCCCGTGA 1062
DB 1112 CAGTATGACAAAATGGAACCTATTAAGATATCCCGTGA 1153
RESULT 5
BD183203
LOCUS BD183203 1062 bp DNA linear PAT 17-JUN-2003
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD183203
VERSION BD183203.1 GI:31875403
KEYWORDS JP 2002345481-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Terao,Y., Matsui,H. and Shintani,Y.
TITLE Novel G protein-coupled receptor and its DNA
JOURNAL Patent: JP 2002345481-A 3 03-DEC-2002;
TAKEDA CHEMICAL INDUSTRIES LTD
OS
FN JP 2002345481-A/3
PD 03-DEC-2002
PF 01-JUN-2001 JP 2001166688
PI YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/09,A61K45/00,A61P1/00,A61P9/00,A61P25/00, PC
A61P29/00,

PC A61P35/00,A61P37/00,C07K14/705,C07K16/28,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC
G01N33/53,
PC G01N33/566,C12N15/00,C12N5/00
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
/organism='Homo sapiens (human)'.
Location/Qualifiers
1..1062
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

source

ORIGIN

Query Match 99.8%; Score 1060.4; DB 6; Length 1062;
Best Local Similarity 99.9%; Pred. No. 1.3e-246;
Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGCAGACGCGCCCACTCGACGCAAGCTCGCTGCTGTGGTGTCCCGGCGC 60
DB 1 ATGAGCAGACGCGCCCACTCGACGCAAGCTCGCTGCTGTGGTGTCCCGGCGC 60
QY 61 TGGGCTGGGGCTTGGGTTTCGTCGCGGTCTACTACAGCTCTTGTGCTGCTCGGT 120
DB 61 TGGGCTGGGGCTTGGGTTTCGTCGCGGTCTACTACAGCTCTTGTGCTGCTCGGT 120
QY 121 TTACAGCAAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAAG 180
DB 121 TTACAGCAAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAAG 180
QY 181 TCTCTTCAAGATATCTTGGCACTGTGCTGTCGCGACATCTTGTCTCTTTTCATA 240
DB 181 TCTCTTCAAGATATCTTGGCACTGTGCTGTCGCGACATCTTGTCTCTTTTCATA 240
QY 241 GGTGTGTGAGCTTCCGTTGGAAATTCATCTTGAAGATGAGATGCTCCAGGTCCC 300
DB 241 GGTGTGTGAGCTTCCGTTGGAAATTCATCTTGAAGATGAGATGCTCCAGGTCCC 300
QY 301 GACCAAGATCATAGAGTGTGATTTCTCATCCATCCACACTCATATGATTTACTGTA 360
DB 301 GACCAAGATCATAGAGTGTGATTTCTCATCCATCCACACTCATATGATTTACTGTA 360
QY 361 CCGTTAACCATTTGACAGATATGCTGTCTGTCACCGCTCAAGTACCAACAGGTCTCA 420
DB 361 CCGTTAACCATTTGACAGATATGCTGTCTGTCACCGCTCAAGTACCAACAGGTCTCA 420
QY 421 TACCCAGCCCGGACCCGGAAAGTCAATTTGATTAAGTTTAACTCACTGCTTCCGACAGC 480
DB 421 TACCCAGCCCGGACCCGGAAAGTCAATTTGATTAAGTTTAACTCACTGCTTCCGACAGC 480
QY 481 ATCCCTATTTATCTGTGAGCCCAATCTGGAAGTGAAGTCACTGAGCACTCTGTGCA 540
DB 481 ATCCCTATTTATCTGTGAGCCCAATCTGGAAGTGAAGTCACTGAGCACTCTGTGCA 540
QY 541 CAGGTCTCATCTGATTCACCTGCTTCAACGCTTACCTGTGCTGCTCATCTTCTTC 600
DB 541 CAGGTCTCATCTGATTCACCTGCTTCAACGCTTACCTGTGCTGCTCATCTTCTTC 600
QY 501 ATTCTTAATCATCATTTGTTGATCAAGCTCAGAGAGAAAGCAATTTTGTCTCCGTGCG 660
DB 501 ATTCTTAATCATCATTTGTTGATCAAGCTCAGAGAGAAAGCAATTTTGTCTCCGTGCG 660
QY 601 ATCTTAATCATCATTTGTTGATCAAGCTCAGAGAGAAAGCAATTTTGTCTCCGTGCG 660
DB 601 ATCTTAATCATCATTTGTTGATCAAGCTCAGAGAGAAAGCAATTTTGTCTCCGTGCG 660
QY 661 TACTCAGAGGAGAGACACCGCATCTTGTGATCACTTACATCTTGTGCAACATTT 720
DB 661 TACTCAGAGGAGAGACACCGCATCTTGTGATCACTTACATCTTGTGCAACATTT 720
QY 721 TGGGCTGGGGCTTGGGTTTCGTCGCGGTCTACTACAGCTCTTGTGCTGCTCGGT 120
DB 721 TGGGCTGGGGCTTGGGTTTCGTCGCGGTCTACTACAGCTCTTGTGCTGCTCGGT 120
QY 781 TGGCTGTGAGATCATGTCCGACATTTGCCAAGATGCTAGCCCTTGTGAACAGAGCATC 840
DB 781 TGGCTGTGAGATCATGTCCGACATTTGCCAAGATGCTAGCCCTTGTGAACAGAGCATC 840

```

```

DB 781 TGGCTGTGAGATCATGTCCGACATTTGCCAAGATGCTAGCCCTTGTGAACAGAGCATC 840
QY 841 AACTTCTTCTTACTGCTTTCATCAGCAAGCGGTTCCGACCAATGGCAGCCGCGACGTC 900
DB 841 AACTTCTTCTTACTGCTTTCATCAGCAAGCGGTTCCGACCAATGGCAGCCGCGACGTC 900
QY 901 AAGGCTTTTTCAGAGCCAGAGCAAGCACTGTACAGTTTACCAATATTAATCTTTC 960
DB 901 AAGGCTTTTTCAGAGCCAGAGCAAGCACTGTACAGTTTACCAATATTAATCTTTC 960
QY 961 ATPACAGATGAGCCCGTGAATCTGCGCGCAACTCAGCTGATCAAGATGCTGTATC 1020
DB 961 ATPACAGATGAGCCCGTGAATCTGCGCGCAACTCAGCTGATCAAGATGCTGTATC 1020
QY 1021 CAGTATGACAAATAATGAAATCTATTAAGATATCCCGGTGA 1062
DB 1021 CAGTATGACAAATAATGAAATCTATTAAGATATCCCGGTGA 1062

```

RESULT 6
BD103851 1062 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD103851
VERSION BD103851.1 GI:22649425
KEYWORDS WO 0194582-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1062)
Terao,Y., Matsumi,H. and Shintani,Y.
Novel G protein-coupled receptor and its DNA
Patent: WO 0194582-A 3 13-DEC-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI
SHINTANI

COMMENT

OS Homo sapiens (human)
PN WO 0194582-A/3
PD 13-DEC-2001
PF 01-JUN-2001 WO 2001JP004643
PR 02-JUN-2000 JP 00P 170446,23-JUN-2000 JP 00P 194926 PI
PC C12N15/12,C12P21/02,C07K14/705,C07K16/28,A61K45/00,A61P25/00,
PC A61P25/00,
PC A61P9/00,A61P35/00,A61P37/00,A61P1/00,G01N33/566, PC
G01N33/50,
PC G01N33/15
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
/organism='Homo sapiens (human)'.
Location/Qualifiers
1..1062
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

source

ORIGIN

Query Match 99.8%; Score 1060.4; DB 6; Length 1062;
Best Local Similarity 99.9%; Pred. No. 1.3e-246;
Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGCAGACGCGCCCACTCGACGCAAGCTCGCTGCTGTGGTGTCCCGGCGC 60
DB 1 ATGAGCAGACGCGCCCACTCGACGCAAGCTCGCTGCTGTGGTGTCCCGGCGC 60
QY 61 TGGGCTGGGGCTTGGGTTTCGTCGCGGTCTACTACAGCTCTTGTGCTGCTCGGT 120
DB 61 TGGGCTGGGGCTTGGGTTTCGTCGCGGTCTACTACAGCTCTTGTGCTGCTCGGT 120
QY 121 TTACAGCAAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAAG 180
DB 121 TTACAGCAAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAAG 180

```

Db	121	TTACGACGAAATATCTTGACAGTATCATCTCTCCAGCTGTTGGCAAGACAGAAAG	180
QY	181	TCCTCTCAACAACATATCTCTTGAGCACTGCGTGCAGCAATCTTTGGCTCTTTTTCATA	240
Db	181	TCCTCTCAACAATATCTCTTGAGCACTGCGTGCAGCAATCTTTGGCTCTTTTTCATA	240
QY	241	GTGTTTGTGACTTCCCTGTTGGAAAGATTTGATCTTGAAACATGACAGATGCTCCAGTCCCC	300
Db	241	GTGTTTGTGACTTCCCTGTTGGAAAGATTTGATCTTGAAACATGACAGATGCTCCAGTCCCC	300
QY	301	GACAGATCATAGAGAGTCTGGAATTCATTCATCCACATCCCTCCATATATGATTAATCTGA	360
Db	301	GACAGATCATAGAGAGTCTGGAATTCATTCATCCACATCCCTCCATATATGATTAATCTGA	360
QY	361	CCGTAAACCATTTGACAGATATATCGCTGTGTGCAACCCGCTCAAGTACCAACAGTCTCA	420
Db	361	CCGTAAACCATTTGACAGATATATCGCTGTGTGCAACCCGCTCAAGTACCAACAGTCTCA	420
QY	421	TACCCAGCCCGCACCCGAAAGCATTTGTAAGTTTATATCATCCTGTTCTGTACACAGC	480
Db	421	TACCCAGCCCGCACCCGAAAGCATTTGTAAGTTTATATCATCCTGTTCTGTACACAGC	480
QY	481	ATCCCTTATTACTGTGTGCCCAACATGTGACTGAAGACTAATAGACACTCTGTGCAT	540
Db	481	ATCCCTTATTACTGTGTGCCCAACATGTGACTGAAGACTAATAGACACTCTGTGCAT	540
QY	541	CACGTCTCATCTGGAATCAGTGTCAACCGTCTACCCGTGACCTGCTGCATCTTCTTC	600
Db	541	CACGTCTCATCTGGAATCAGTGTCTACCGTCTACCCGTGACCTGCTGCATCTTCTTC	600
QY	601	ATCTTGAATCAATCATCTTGTGTACAAGCTCAGAGAGAAAGCAATTTTGTCTCCGTGC	660
Db	601	ATCTTGAATCAATCATCTTGTGTACAAGCTCAGAGAGAAAGCAATTTTGTCTCCGTGC	660
QY	661	TACTCCACGGGGAGAACACCGCATCTTGTTCACATTAAGCTCATCTTGTGCAACATTT	720
Db	661	TACTCCACGGGGAGAACACCGCATCTTGTTCACATTAAGCTCATCTTGTGCAACATTT	720
QY	721	TGGGCCCCCGCATCATATGATTTCTTTACCACTCTTATGGGGCGCCCATCGAAGACCG	780
Db	721	TGGGCCCCCGCATCATATGATTTCTTTACCACTCTTATGGGGCGCCCATCGAAGACCG	780
QY	781	TGGCTGTGTGACATCATGTGCCGACATTTGCCAATCATGTAGCCCTTTGAAACAGCAATC	840
Db	781	TGGCTGTGTGACATCATGTGCCGACATTTGCCAATCATGTAGCCCTTTGAAACAGCAATC	840
QY	841	AACTTCTTCTCTACTGCTTCATCAGCAAGGGGTTCGCAACATGAGAGCGCCACGCTC	900
Db	841	AACTTCTTCTCTACTGCTTCATCAGCAAGGGGTTCGCAACATGAGAGCGCCACGCTC	900
QY	901	AAGGCTTCTCAAGTGCAGAAACCTGTATACAGTTCTACACCAATCATATATCTTTCC	960
Db	901	AAGGCTTCTCAAGTGCAGAAACCTGTGTACAGTTCTACACCAATCATATCTTTTC	960
QY	961	ATTAACAAGTAGCCCCCTGATCTCGCCGGCAAACTCAACATGCAATCAAGAGTCTGTGTAC	1020
Db	961	ATTAACAAGTAGCCCCCTGATCTCGCCGGCAAACTCAACATGCAATCAAGAGTCTGTGTAC	1020
QY	1021	CAGATATGCAAAAAATGAAAAACCTATATAAAGTATCCCGGTGA	1062
Db	1021	CAGATATGCAAAAAATGAAAAACCTATATAAAGTATCCCGGTGA	1062

RESULT	7			
LOCUS	AX254977			
DEFINITION	Sequence 21 from Patent WO01/0978.	1343 bp	DNA	linear
ACCESSION	AX254977			PAT 10-OCT-2001
VERSION	AX254977.1	GI:16074504		
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
1	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	Taupier, R.J., Majumder, K., Spaderna, S.K., Smithson, G., Mezes, P.S. and Vernot, C.A.	Polypeptides and nucleic acids encoding same	Patent: WO 0170978-A 21 27-SEP-2001; Curagen Corporation (US)	location/Qualifiers 1..1143 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"

Query Match	99.7%	Score 1058.8;	DB 6;	Length 1343;
Best Local Similarity	99.8%	Pred. No. 3.2e-246;		
Matches 1060; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	1	ATGAGCACA	CAGCAGCCACCCTCGAGCCAA	CAGCTCGCTCTTGGTGGTCCCGGAC	60
Db	2	ATGAGACA	CAGCAGCCACCCTCGAGCCAA	CAGCTCGCTCTTGGTGGTCCCGGAC	61
QY	61	TGGGCTGGGGCTTGGGTTTCTGGCCCTGGTCTA	CAAGCTCTTGCTGGCTCGT	120	
Db	62	TGGGCTGGGGCTTGGGTTTCTGGCCCTGGTCTA	CAAGCTCTTGCTGGCTCGT	121	
QY	121	TTACGCAAAATATCTTGCACATGATATCTCTCCGAC	TGGTGGGAGAGACAGAG	180	
Db	122	TTACGCAAAATATCTTGCACATGATATCTCTCCGAC	TGGTGGGAGAGACAGAG	181	
QY	181	TCTCTTACATATCTCTTGGCACTGCTCTGCGACAT	CTTGTCCTCTTTTTCATA	240	
Db	182	TCTCTCTTACATATCTCTTGGCACTGCTCTGCGACAT	CTTGTCCTCTTTTTCATA	241	
QY	241	GTTGTTGGAGACTTCTCTGTGGAGATTCACTTGAAC	ATGCAATGCTTCAGCTCCC	300	
Db	242	GTTGTTGGAGACTTCTCTGTGGAGATTCACTTGAAC	ATGCAATGCTTCAGCTCCC	301	
QY	301	GACAAATCATAGAGTGTGAAATTCATCATCACA	CACTTCATATGATTAATCTGA	360	
Db	302	GACAAATCATAGAGTGTGAAATTCATCATCACA	CACTTCATATGATTAATCTGA	361	
QY	361	CGGTAAACATTTGACAGGTATATGCTGTCTGCCAC	CCGCTGAAATACACACGGTCTCA	420	
Db	362	CGGTAAACATTTGACAGGTATATGCTGTCTGCCAC	CCGCTGAAATACACACGGTCTCA	421	
QY	421	TACCAAGCCGCAACCCGGAAAGCATTTGAAGTATT	TATACATCACTCTCCGACACAG	480	
Db	422	TACCAAGCCGCAACCCGGAAAGCATTTGAAGTATT	TATACATCACTCTCCGACACAG	481	
QY	481	ATCCCTTATATCTGTGGCCCAACATCTGGAATGAA	CACTATACACACTCTGTGCAT	540	
Db	482	ATCCCTTATATCTGTGGCCCAACATCTGGAATGAA	CACTATACACACTCTGTGCAT	541	
QY	541	CAGCTCTCATATCTGATTCACCTGTTCAACGCTTAC	CTGGTGGCCCTGCTTCATCTTCTTC	600	
Db	542	CAGCTCTCATATCTGATTCACCTGTTCAACGCTTAC	CTGGTGGCCCTGCTTCATCTTCTTC	601	
QY	601	ATCTTGAATCAATCATTTGTGTCAAGACTCAGAGAG	AAGACAATTTGTCCTCGTGC	660	
Db	602	ATCTTGAATCAATCATTTGTGTGTCAAGAGCTCAG	AGAGACAATTTGTCCTCGTGC	661	
QY	661	TACTCCACGGGAGAGACACACCGCATCTTGTGTCA	CACTTAACCTCCATCTTGGCACACTT	720	
Db	662	TACTCCACGGGAGAGAGACACACCGCATCTTGTGT	CACTTAACCTCCATCTTGGCACACTT	721	
QY	721	TGGGCCCCCGCATCATATGATTTCTTTTACACCT	TATGGGGGCGCATTCAGAACCGC	780	
Db	722	TGGGCCCCCGCATCATATGATTTCTTTTACACCT	TATGGGGGCGCATTCAGAACCGC	781	
QY	781	TGGCTGTGTGACATCATGTGCCACATTCGCAAT	GCTTGTGAAACACAGCCATTC	840	
Db	782	TGGCTGTGTGACATCATGTGCCACATTCGCAAT	GCTTGTGAAACACAGCCATTC	841	

QY 841 AACTGCTCTCTCTACTGCTTCATCAGCAAGGGTTCGGACCATGAGAGCGCCAGCCTC 900
DB 842 AACTCTCTCTCTCTACTGCTTCATCAGCAAGGGTTCGGACCATGAGAGCGCCAGCCTC 901
QY 901 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTACACCAATCAATCTTTCC 960
DB 902 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTACACCAATCAATCTTTCC 961
QY 961 ATACAGATGAGCCCTGATCTTCGGCGGCAACTCACTGCATCAAGATGCTGTGTAC 1020
DB 962 ATACAGATGAGCCCTGATCTTCGGCGGCAACTCACTGCATCAAGATGCTGTGTAC 1021
QY 1021 CAGTATGACAAATAAGTAAATCTATAAAGTATCCCCGTGA 1062
DB 1022 CAGTATGACAAATAAGTAAATCTATAAAGTATCCCCGTGA 1063

RESULT 8
AX463227 1059 bp DNA linear PAT 15-JUL-2002
LOCUS AX463227 Sequence 1 from Patent WO248358.
DEFINITION AX463227
ACCESSION AX463227
VERSION AX463227.1 GI:21886199
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Smolyar, A., Zhu, Z., Encinas, J., Watanabe, S. and Okigami, H.
AUTHORS Regulation of human chemokine-like receptor
TITLE Patent: WO 0248358-A 1 20-JUN-2002;
JOURNAL Bayer Aktiengesellschaft (DE)
FEATURES
source 1..1059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.6%; Score 1057.4; DB 6; Length 1059;
Best Local Similarity 99.9%; Pred. No. 6.8e-246;
Matches 1058; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACACGACGCGCCACTCGACGCCAAGCTCGCTGTTGGTGGTCCCGGCG 60
DB 1 ATGAGACACACGACGCGCCACTCGACGCCAAGCTCGCTGTTGGTGGTCCCGGCG 60
QY 61 TCGGCTGGGCTGGGCTTTCGTCGCGCGGTCTTACTACAGCTTTGCTGCTCGCT 120
DB 61 TCGGCTGGGCTGGGCTTTCGTCGCGCGGTCTTACTACAGCTTTGCTGCTCGCT 120
QY 121 TTACACGACAAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGAGAGAG 180
DB 121 TTACACGACAAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGAGAGAG 180
QY 181 TCTCTCTCAACTATCTTTGGAGCTCGCTGCGGACACTTTGGTCTCTTTTTCATA 240
DB 181 TCTCTCTCAACTATCTTTGGAGCTCGCTGCGGACACTTTGGTCTCTTTTTCATA 240
QY 241 GTGTTTGTGAGCTCTCTGTTTGAAGATTTCATTTGAACATGAGATGCTCGCTG 300
DB 241 GTGTTTGTGAGCTCTCTGTTTGAAGATTTCATTTGAACATGAGATGCTCGCTG 300
QY 301 GACAAATATGAGATGCTGAGATTTCTCATCTCCACACTCCATATGATTAATCTGTA 360
DB 301 GACAAATATGAGATGCTGAGATTTCTCATCTCCACACTCCATATGATTAATCTGTA 360
QY 361 CCGTTAACATTTGACAGATATGCTGTCTGCGACCGGCTCAAGTACACACGCTTCA 420
DB 361 CCGTTAACATTTGACAGATATGCTGTCTGCGACCGGCTCAAGTACACACGCTTCA 420

QY 421 TACCCAGCCCGACCCGAAAGATGTAAGTGTATATCATCACTGCTTCTGACGAC 480
DB 421 TACCCAGCCCGACCCGAAAGATGTAAGTGTATATCATCACTGCTTCTGACGAC 480
QY 481 ATCCCTATATTACTGATGAGCCCAACATCTGGAATGAAGATCAATCAGCACTTGTGAT 540
DB 481 ATCCCTATATTACTGATGAGCCCAACATCTGGAATGAAGATCAATCAGCACTTGTGAT 540
QY 541 CAGTCTCATCTGATATCCATGCTTCAACCGGTCTACCTGCTGCTGCTTCTTTC 600
DB 541 CAGTCTCATCTGATATCCATGCTTCAACCGGTCTACCTGCTGCTGCTTCTTTC 600
QY 601 ATCTGAATCATATCTTGTATGACAGCTCAGAGAGAGAGCAATTTTGTCTCGTGGC 660
DB 601 ATCTGAATCATATCTTGTATGACAGCTCAGAGAGAGAGCAATTTTGTCTCGTGGC 660
QY 661 TACTCCAGCGGAGAGACACCGCATCTTGTTCACCATTAACCTCATCTTTGACACTT 720
DB 661 TACTCCAGCGGAGAGACACCGCATCTTGTTCACCATTAACCTCATCTTTGACACTT 720
QY 721 TGGGCCCCCGGACATCATCATGATTTTTCACCTTATGAGGCGCCCATCCAGAACCG 780
DB 721 TGGGCCCCCGGACATCATCATGATTTTTCACCTTATGAGGCGCCCATCCAGAACCG 780
QY 781 TGGCTGTGACATCATGATTCGACCATGACCATGCTGAGCCCTTGAACACAGCATC 840
DB 781 TGGCTGTGACATCATGATTCGACCATGACCATGCTGAGCCCTTGAACACAGCATC 840
QY 841 AACTCTTCTCTACTGCTTCTCATCAGCAAGCGGTTCCGACCATGAGCGCGCAGCTC 900
DB 841 AACTCTTCTCTACTGCTTCTCATCAGCAAGCGGTTCCGACCATGAGCGCGCAGCTC 900
QY 901 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTACACCAATCAATCTTTCC 960
DB 901 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTACACCAATCAATCTTTCC 960
QY 961 ATACAGATGAGCCCTGATCTTCGGCGGCAACTCACTGCATCAAGATGCTGTGTAC 1020
DB 961 ATACAGATGAGCCCTGATCTTCGGCGGCAACTCACTGCATCAAGATGCTGTGTAC 1020
QY 1021 CAGTATGACAAATAAGTAAATCTATAAAGTATCCCCG 1059
DB 1021 CAGTATGACAAATAAGTAAATCTATAAAGTATCCCCG 1059

RESULT 9
AX664703 1062 bp DNA linear PAT 22-MAR-2003
LOCUS AX664703 Sequence 12 from Patent WO0207960.
DEFINITION AX664703
ACCESSION AX664703
VERSION AX664703.1 GI:29164463
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Leiby, K.R., Kapeller-Libermann, R. and Glucksmann, M.
AUTHORS 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
TITLE of human proteins and uses thereof
JOURNAL Patent: WO 02074960-A 12-26-SEP-2002;
Milleium Pharmaceuticals, Inc. (US)
FEATURES
source 1..1062
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.9%; Score 1050.8; DB 6; Length 1062;
Best Local Similarity 99.3%; Pred. No. 2.7e-244;
Matches 1055; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 1 ATGAGACACAGCAGCCGACCTGGAGACCAAGCTGCTGTGTGGTGGTCCCCGGC 60
Db 1 ATGAGACACAGCAGCCGACCTGGAGACCAAGCTGCTGTGTGGTGGTCCCCGGC 60
QY 61 TCGGCTCGGGCTGGGTTTGTGGCCGTGGTCTAATACAGCCTCTTGCTGGCTGGT 120
Db 61 TCGGCTCGGGCTGGGTTTGTGGCCGTGGTCTAATACAGCCTCTTGCTGGCTGGT 120
QY 121 TTACAGCAAAATCTTGAAGTATATCTCTCCAGCTGGTGGCAGAGAGAGAG 180
Db 121 TTACAGCAAAATCTTGAAGTATATCTCTCCAGCTGGTGGCAGAGAGAGAG 180
QY 181 TCTCTCAAACTATCTTGTGGCACTGGCTGGCCAGACATCTTGCTCTTTTATA 240
Db 181 TCTCTCAAACTATCTTGTGGCACTGGCTGGCCAGACATCTTGCTCTTTTATA 240
QY 241 GTGTTTGGACCTTCTGTTGGAAATTTCACTTGAACATGAGATGCTCCAGTCCC 300
Db 241 GTGTTTGGACCTTCTGTTGGAAATTTCACTTGAACATGAGATGCTCCAGTCCC 300
QY 301 GACAGATCATAGAAAGTGGAAATTCATCCATCCACACTCCATATGATTAATCTGA 360
Db 301 GACAGATCATAGAAAGTGGAAATTCATCCATCCACACTCCATATGATTAATCTGA 360
QY 361 CCGTTAACATTTGACAGATATATGCTGTGGCAGCCGCTCAAGTACACAGGTCTGA 420
Db 361 CCGTTAACATTTGACAGATATATGCTGTGGCAGCCGCTCAAGTACACAGGTCTGA 420
QY 421 TACCCAGCCGACCCGGAAGTCAATTTGAAGTGTATATCACTGCTTCTGACACAG 480
Db 421 TACCCAGCCGACCCGGAAGTCAATTTGAAGTGTATATCACTGCTTCTGACACAG 480
QY 481 ATCCCTATATAGTGTGGCCCAACATCTGAGCTGAAGACATCACTGAGCCTGTGGAT 540
Db 481 ATCCCTATATAGTGTGGCCCAACATCTGAGCTGAAGACATCACTGAGCCTGTGGAT 540
QY 541 CAGCTCTCATCTGATATCCATCTCTTCACTGCTTCACTGCTGCTGCTCAATCTTCTTC 600
Db 541 CAGCTCTCATCTGATATCCATCTCTTCACTGCTTCACTGCTGCTGCTCAATCTTCTTC 600
QY 601 ATCTTGAATCATCATTTGTGTATCAAGCTCAAGAGAGAGAGAAATTTGTCTGTGGC 660
Db 601 ATCTTGAATCATCATTTGTGTGTATCAAGCTCAAGAGAGAGAGAAATTTGTCTGTGGC 660
QY 661 TACTCAGCGGAGAGACACCGCATCTTGTTCACCATTAATCTCAATCTTGTGCAACAT 720
Db 661 TACTCAGCGGAGAGACACCGCATCTTGTTCACCATTAATCTCAATCTTGTGCAACAT 720
QY 721 TGGGCCCCCGCATCATATGATTTCTTACACACTCTTATGAGGCGCCCATCCAGAACGC 780
Db 721 TGGGCCCCCGCATCATATGATTTCTTACACACTCTTATGAGGCGCCCATCCAGAACGC 780
QY 781 TGGCTGTGACATATGTCGACATTTGCAATGCAATGCTAGCCCTCTGAGACAGCATC 840
Db 781 TGGCTGTGACATATGTCGACATTTGCAATGCAATGCTAGCCCTCTGAGACAGCATC 840
QY 841 AACTCTCTCTTACTGCTTATCAAGAGCGGTTCCGACCATGAGCCGCAAGCTC 900
Db 841 AACTCTCTCTTACTGCTTATCAAGAGCGGTTCCGACCATGAGCCGCAAGCTC 900
QY 901 AAGGCTTCTTCAATGTCAGAGAACCTGTGACAGTTCTAACAACATTAATCTTTC 960
Db 901 AAGGCTTCTTCAATGTCAGAGAACCTGTGACAGTTCTAACAACATTAATCTTTC 960
QY 961 ATAAAGATGAGCCCTGGATCTCGCGGCAAACTCACACTGATCAAGATCTGTGTAC 1020
Db 961 ATAAAGATGAGCCCTGGATCTCGCGGCAAACTCACACTGATCAAGATCTGTGTAC 1020
QY 1021 CAGTATGACAAATATGAAAACCTTAAAGTATCCCGCTGA 1062
Db 1021 CAGTATGACAAATATGAAAACCTTAAAGTATCCCGCTGA 1062
```

```
RESULT 10
AX664701 1526 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 10 from Patent WO02074960.
DEFINITION AX664701
ACCESSION AX664701
VERSION AX664701.1 GI:29164461
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butioria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
Leidy, K.R., Kapeller-Liebermann, R. and Glucksman, M.
38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
of human proteins and uses thereof
Patent: WO 02074960-A 10 26-SEP-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..1526
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
139..1200
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD80046.1"
/db_xref="GI:29164462"
/translation="MEHTHAILANSLSLWMPGACGLGFPVYVYSLIQCIGLPPN
ILVILISQVARKSSNYLALADLVLFYVDFLEDPILNMQOVPK
ILEVBSISHTSITWVPLTIDRYIAVCHPLKHTYISFARKYIVSYITCFIS
IPYWMNIMEDIDISTVHVHVLWCFIVYDPSIFILNSIVYKLRNSFEL
CGYSTGKTALIFLTISIFAILWAPRILIMLYHYPIONRLVHIMSDIANNMLL
NTAINFLYFISIRFRFMAATLKAFKOKOPVFPYTHNFSITSSPWSANSHC
IKMLVYDNGKRIKVS"
ORIGIN
Query Match 98.9%; Score 1050.8; DB 6; Length 1526;
Best Local Similarity 99.3%; Pred. No. 2.6e-244;
Matches 1051; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATGAGACACAGCAGCCGACCTGGAGACCAAGCTGCTGTGTGGTGGTCCCCGGC 60
Db 139 ATGAGACACAGCAGCCGACCTGGAGACCAAGCTGCTGTGTGGTGGTCCCCGGC 198
QY 61 TCGGCTCGGGCTGGGTTTGTGGCCGTGGTCTAATACAGCCTCTTGCTGGCTGGT 120
Db 199 TCGGCTCGGGCTGGGTTTGTGGCCGTGGTCTAATACAGCCTCTTGCTGGCTGGT 258
QY 121 TTACAGCAAAATATCTTGAAGTATATCTCTCCAGCTGGTGGCAGAGAGAG 180
Db 259 TTACAGCAAAATATCTTGAAGTATATCTCTCCAGCTGGTGGCAGAGAGAG 318
QY 181 TCTCTCAAACTATCTTGTGGCACTGGCTGGCCAGACATCTTGCTCTTTTATA 240
Db 319 TCTCTCAAACTATCTTGTGGCACTGGCTGGCCAGACATCTTGCTCTTTTATA 378
QY 241 GTGTTTGGACCTTCTGTTGGAAATTTCACTTGAACATGAGATGCTCCAGTCCC 300
Db 379 GTGTTTGGACCTTCTGTTGGAAATTTCACTTGAACATGAGATGCTCCAGTCCC 438
QY 301 GACAGATCATAGAAAGTGGAAATTCATCCATCCACACTCCATATGATTAATCTGA 360
Db 439 GACAGATCATAGAAAGTGGAAATTCATCCATCCACACTCCATATGATTAATCTGA 498
QY 361 CCGTTAACATTTGACAGATATATGCTGTGGCAGCCGCTCAAGTACACAGGTCTGA 420
Db 499 CCGTTAACATTTGACAGATATATGCTGTGGCAGCCGCTCAAGTACACAGGTCTGA 558
QY 421 TACCCAGCCGACCCGGAAGTCAATTTGAAGTGTATATCACTGCTTCTGACACAG 480
Db 559 TACCCAGCCGACCCGGAAGTCAATTTGAAGTGTATATCACTGCTTCTGACACAG 618
QY 481 ATCCCTATATAGTGTGGCCCAACATCTGAGCTGAAGACATCACTGAGCCTGTGGAT 540
```


Db 619 ATCCCATATGAGTGGGCCCAACATGGAAGTGAACATCACTACAGCACTCTGTGAT 678
QY 541 CAGCTCTCATCTGGATTCACCTGTTCAACGCTTACCTGGTCCCTGCTCATCTTTCTT 600
Db 679 CAGCTCTCATCTGGGTCACCTGCTTCAACGCTTACAGAGTGCCTGCTCATCTTTCTT 728
QY 601 ATCTTAACATCAATCTTGTGTACAAAGCTCAGAGAGAGAAATTTTGTCTTCCGTGGC 660
Db 739 ATCTTAACATCAATCTTGTGTACAAAGCTCAGAGAGAGAAATTTTGTCTTCCGTGGC 798
QY 661 TACTCAGGGGGAAGACACACCGCATCTTGTTCACCACTTACCTCATCTTTGCGACACT 720
Db 799 TACTCAGGGGGAAGACACACCGCATCTTGTTCACCACTTACCTCATCTTTGCGACACT 858
QY 721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTTATGAGGCGCCCATCAGAACCGC 780
Db 859 TGGGCCCCCGCATCATCATGATTTCTTTACCACTTATGAGGCGCCCATCAGAACCGC 918
QY 781 TGGCTGTGACATCATATGTCCTCCGACATTTGCCAATGCTAGCCTTTCTGAAACAGCCATC 840
Db 919 TGGCTGTGACATCATATGTCCTCCGACATTTGCCAATGCTAGCCTTTCTGAAACAGCCATC 978
QY 841 AACTTCTCTCTACTGCTTCTCATCAGCAGCGGTTCCGACACATGCGAGCGCCAGCCTC 900
Db 979 AACTTCTCTCTACTGCTTCTCATCAGCAGCGGTTCCGACACATGCGAGCGCCAGCCTC 1038
QY 901 AAGGCTTTCTTCAAGTCCAGAAAGCAACCTGTACAGTTCTACACCAATCAATCTTTTCC 960
Db 1039 AAGGCTTTCTTCAAGTCCAGAAAGCAACCTGTACAGTTCTACACCAATCAATCTTTTCC 1098
QY 961 ATTAACAAGTACGCCCCCTGGATCTGCGCGGCAACTCAGCTGCTCAAGAGTCTGTGTAC 1020
Db 1099 ATTAACAAGTACGCCCCCTGGATCTGCGCGGCAACTCAGCTGCTCAAGAGTCTGTGTAC 1158
QY 1021 CAGTATGACAAAATGAAAACTATAAAGTATCCCGGTGA 1062
Db 1159 CAGTATGACAAAATGAAAACTATAAAGTATCCCGGTGA 1200

RESULT 11
AX254975 1119 bp DNA linear PAT 10-OCT-2001
LOCUS AX254975
DEFINITION Sequence 19 from Patent WO0170978.
ACCESSION AX254975
VERSION AX254975.1 GI:16074503
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catartida; Homnidae; Homo.
REFERENCE 1
AUTHORS Taupier, R.J., Majumder, K., Spaderna, S.K., Smithson, G., Mezes, P.S.
and Vernet, C.A.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0170978-A 19 27-SEP-2001;
FEATURES
source
1. 1119
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.7%; Score 1048.4; DB 6; length 1119;
Best Local Similarity 99.9%; Pred. No. 1e-243;
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACACGACGCCCACTTCGACGCCAAGAGCTGCTGTGTGTGTCCTCCCGGC 60
Db 1 ATGAGACACACGACGCCCACTTCGACGCCAAGAGCTGCTGTGTGTGTCCTCCCGGC 60
QY 61 TCGCGCTGCGGCTTGGGTTTGTGCGCGGTGTCTAATCAAGCTCTTGTGTGCTCGGT 120

Db 61 TCGCGCTGCGGCTTGGGTTTGTGCGCGGTGTCTAATCAAGCTCTTGTGTGCTCGGT 120
QY 121 TTACAGCAAAATCTTGAACAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAAG 180
Db 121 TTACAGCAAAATCTTGAACAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAAG 180
QY 181 TCTCTTACATATCTTGTGCACTGCTGCTGCGGACATCTTGGTCTCTTTTTCATA 240
Db 181 TCTCTTACATATCTTGTGCACTGCTGCTGCGGACATCTTGGTCTCTTTTTCATA 240
QY 241 GGTGTGTGAACTTCTGTGGAAATTCATCTTGAACATGAGATGCTCAGGTCGCC 300
Db 241 GGTGTGTGAACTTCTGTGGAAATTCATCTTGAACATGAGATGCTCAGGTCGCC 300
QY 301 GACAAATCATAGAAAGTGTGAATTTCTATCATCAGACCTCATATGATTAATGTA 360
Db 301 GACAAATCATAGAAAGTGTGAATTTCTATCATCAGACCTCATATGATTAATGTA 360
QY 361 CGGTAAACATTAACAGATATATCGGTGTGCAACCCGCTCAAGTACACAGGTCTCA 420
Db 361 CGGTAAACATTAACAGATATATCGGTGTGCAACCCGCTCAAGTACACAGGTCTCA 420
QY 421 TACCCAGCCCGACCCGGAAGATGTAAGTGTTAATCATCACTGCTTCTGACAGC 480
Db 421 TACCCAGCCCGACCCGGAAGATGTAAGTGTTAATCATCACTGCTTCTGACAGC 480
QY 481 ATCCCTTAATTAAGTGTGACCAACATCTGGAATGAACTAATGACACCTGTGAT 540
Db 481 ATCCCTTAATTAAGTGTGACCAACATCTGGAATGAACTAATGACACCTGTGAT 540
QY 541 CAGTCTCTATCTGATTCACATGCTTCAACGCTTACCTGCTGCTGCTCTCTTCTTC 600
Db 541 CAGTCTCTATCTGATTCACATGCTTCAACGCTTACCTGCTGCTGCTCTCTTCTTC 600
QY 601 ATCTTAACATCAATCTTGTGTAAGAGCTCAGAGAGAGAGCAATTTTGTCTCCGTGGC 660
Db 601 ATCTTAACATCAATCTTGTGTAAGAGCTCAGAGAGAGAGCAATTTTGTCTCCGTGGC 660
QY 661 TACTCAGGGGGAAGACACACCGCATCTTGTTCACCATTAACCTCTTGTGCAACACT 720
Db 661 TACTCAGGGGGAAGACACACCGCATCTTGTTCACCATTAACCTCTTGTGCAACACT 720
QY 721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTTATGAGGCGCCCATCAGAACCGC 780
Db 721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTTATGAGGCGCCCATCAGAACCGC 780
QY 781 TGGCTGTGACATCATATGTCCTCCGACATTTGCCAATGCTAGCCTTTCTGAAACAGCCATC 840
Db 781 TGGCTGTGACATCATATGTCCTCCGACATTTGCCAATGCTAGCCTTTCTGAAACAGCCATC 840
QY 841 AACTTCTCTCTACTGCTTCTCATCAGCAGCGGTTCCGACCATGCGACCGCAGCTC 900
Db 841 AACTTCTCTCTACTGCTTCTCATCAGCAGCGGTTCCGACCATGCGACCGCAGCTC 900
QY 901 AAGGCTTTCTTCAAGTCCAGAAAGCAACCTGTACAGTTCTTACACCAATATACTTTTC 960
Db 901 AAGGCTTTCTTCAAGTCCAGAAAGCAACCTGTACAGTTCTTACACCAATATACTTTTC 960
QY 961 ATTAACAAGTACGCCCCCTGGATCTGCGCGGCAACTCAGCTGATCAAGATGCTGTGTAC 1020
Db 961 ATTAACAAGTACGCCCCCTGGATCTGCGCGGCAACTCAGCTGATCAAGATGCTGTGTAC 1020
QY 1021 CAGTATGACAAAATGAAAACTATAA 1050
Db 1021 CAGTATGACAAAATGAAAACTATAA 1050

RESULT 12
AX375235 1130 bp DNA linear PAT 01-MAR-2002
LOCUS AX375235
DEFINITION Sequence 25 from Patent WO0210387.
ACCESSION AX375235

QY 301 GACAGATCATAGAGTGTGGAATTCATCCATCCACACCTCCATATGATTACTGTA 360
 |||||
 Db 301 GACAGATCATAGAGTGTGGAATTCATCCATCCACACCTCCATATGATTACTGTA 360
 QY 361 CCGTTAACATTGACAGATATATCGCTGTGCGACCCGCTCAAGTACACACAGCTCTCA 420
 |||||
 Db 361 CCGTTAACATTGACAGATATATCGCTGTGCGACCCGCTCAAGTACACACAGCTCTCA 420
 QY 421 TACCCAGCCCGCACCCGGAAGTCATTGTAAGTGTTAATCACTACCTGCTCTGACACG 480
 |||||
 Db 421 TACCCAGCCCGCACCCGGAAGTCATTGTAAGTGTTAATCACTACCTGCTCTGACACG 480
 QY 481 ATCCCTTATTAATCGTGTGCGCAACATCTGGAATGGAAGTACATGACACCTGTGACAT 540
 |||||
 Db 481 ATCCCTTATTAATCGTGTGCGCAACATCTGGAATGGAAGTACATGACACCTGTGACAT 540
 QY 541 CAGTCTCATCTGATCTGATCTGCTTACCGTGTACCTGAGTGGCTGCTCATCTTCTTC 600
 |||||
 Db 541 CAGTCTCATCTGATCTGATCTGCTTACCGTGTACCTGAGTGGCTGCTCATCTTCTTC 600
 QY 601 ATCTGACTCAATCACTGTGTGTAAGCTCAGAGGAAGAGCAATTTTGTCTCGTGC 660
 |||||
 Db 601 ATCTGACTCAATCACTGTGTGTAAGCTCAGAGGAAGAGCAATTTTGTCTCGTGC 660
 QY 661 TACTCAGCGGGAAGACACCGCATCTGTTCACATTAACCTCCATCTTTGCGACACTT 720
 |||||
 Db 661 TACTCAGCGGGAAGACACCGCATCTGTTCACATTAACCTCCATCTTTGCGACACTT 720
 QY 721 TGGGCCCCCGCATCTGATCTGATCTTTCACACCTCTATGGGCGCCATCCAGAACCG 780
 |||||
 Db 721 TGGGCCCCCGCATCTGATCTGATCTTTCACACCTCTATGGGCGCCATCCAGAACCG 780
 QY 781 TGGCTGTGCAATCATGTGTCGACATTTGCCAACATGACCTTCTGAAACAGGCATC 840
 |||||
 Db 781 TGGCTGTGCAATCATGTGTCGACATTTGCCAACATGACCTTCTGAAACAGGCATC 840
 QY 841 AACTTCTCTCTACCTGCTTCACTAGAGCGGTTCGACCATGAGCAGCCGCGACGCTC 900
 |||||
 Db 841 AACTTCTCTCTACCTGCTTCACTAGAGCGGTTCGACCATGAGCAGCCGCGACGCTC 900
 QY 901 AAGCTTTCTTCAAGTGCAGAGAACCACTGTACAGTCTTCAACCAATCAATCTTTTC 960
 |||||
 Db 901 AAGCTTTCTTCAAGTGCAGAGAACCACTGTACAGTCTTCAACCAATCAATCTTTTC 960
 QY 961 ATPAACAGTAGCCCTGTGATCTCGCGGAACTCACTGATCAAGATGCTGTGAC 1020
 |||||
 Db 961 ATPAACAGTAGCCCTGTGATCTCGCGGAACTCACTGATCAAGATGCTGTGAC 1020
 QY 1021 CAGTATGACAAATAAGTAAAGTATCCCGGTGA 1062
 |||||
 Db 1021 CAGTATGACAAATAAGTAAAGTATCCCGGTGA 1062
 RESULT 14
 CQ737667 957 bp DNA linear PAT 03-FEB-2004
 LOCUS CQ737667
 DEFINITION Sequence 23601 from Patent WO02068579.
 ACCESSION CQ737667
 VERSION CQ737667.1 GI:42336565
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 Kite, such as nucleic acid arrays, comprising a majority of
 humerous or transcripts, for detecting expression and other uses
 thereof
 Patent: WO 02068579-A 23601 06-SEP-2002;
 JOURNAL
 PE Corporation (NY) (US)
 Location/Qualifiers
 1. 957
 source

ORIGIN

/organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

Query Match 88.5%; Score 940.2; DB 6; Length 957;
 Best Local Similarity 99.7%; Pred. No. 26-217;
 Matches 942; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 118 GGTTAACGACAAATATCTTGAACAGTATCATCTCTCCAGCTGGTGCAAGAACAG 177
 |||||
 Db 13 GGTTTGACGAAATATCTTGAACAGTATCATCTCTCCAGCTGGTGCAAGAACAG 72
 QY 178 AGTCTCTTCAACATATCTTGGACATCGCTGCTGCCGACATCTTGCTCTTTTC 237
 |||||
 Db 73 AGTCTCTTCAACATATCTTGGACATCGCTGCTGCCGACATCTTGCTCTTTTC 132
 QY 238 ATAGTGTGTGGAATCTCTGTGGAAGATTCATCTTGAACATGACAGTGCAGATC 297
 |||||
 Db 133 ATAGTGTGTGGAATCTCTGTGGAAGATTCATCTTGAACATGACAGTGCAGATC 192
 QY 298 CCGACAGATCATGAAGTGTGGAATTCATCATCATCAGACCTCATATGATTACT 357
 |||||
 Db 193 CCGACAGATCATGAAGTGTGGAATTCATCATCATCAGACCTCATATGATTACT 252
 QY 358 GTACGTTAACATTTGACAGATATCGCTGTGCGACCCGCTCAAGTACACAGGTC 417
 |||||
 Db 253 GTACGTTAACATTTGACAGATATCGCTGTGCGACCCGCTCAAGTACACAGGTC 312
 QY 418 TCATACCCAGCCGGAACCGGGAAGTATGTAAGTGTTAATGACACCTGCTCCGACC 477
 |||||
 Db 313 TCATACCCAGCCGGAACCGGGAAGTATGTAAGTGTTAATGACACCTGCTCCGACC 372
 QY 478 AGCATCCCTTATTAATCTGTGCGCAACATCTGGAATGGAAGTATGATGACAGCCTCTGTG 537
 |||||
 Db 373 AGCATCCCTTATTAATCTGTGCGCAACATCTGGAATGGAAGTATGATGACAGCCTCTGTG 432
 QY 538 CATACGCTCTCATCTGATCTGATCTTACCGCTTACCTGTGCTGCTCATCTTC 597
 |||||
 Db 433 CATACGCTCTCATCTGATCTGATCTTACCGCTTACCTGTGCTGCTCATCTTC 492
 QY 598 TTCACTTGAATCAATCATGTGTGTAAGTCAAGTCAAGAGGAAGCAATTTGTCTCGT 657
 |||||
 Db 493 TTCACTTGAATCAATCATGTGTGTAAGTCAAGTCAAGAGGAAGCAATTTGTCTCGT 552
 QY 658 GGTATCTCAGCGGGAAGACACCGCATCTTGTACCATTAACCTCATCTTTGCCACA 717
 |||||
 Db 553 GGTATCTCAGCGGGAAGACACCGCATCTTGTACCATTAACCTCATCTTTGCCACA 612
 QY 718 CTTTGGGCCCCCGCATCATGATCTTTTACACCTCTATGGGCGCCATCCAGAAC 777
 |||||
 Db 613 CTTTGGGCCCCCGCATCATGATCTTTTACACCTCTATGGGCGCCATCCAGAAC 672
 QY 778 CGTGGCTGTGCAATCATGTGCGACATGCGCAACATGATGCTCTGTAACAGGCC 837
 |||||
 Db 673 CGTGGCTGTGCAATCATGTGCGACATGCGCAACATGATGCTCTGTAACAGGCC 732
 QY 838 ATCAACTTCTTCTCTACTGCTTTCATGACAGAGCGGTTCCGACCATGAGCAGCCGACG 897
 |||||
 Db 733 ATCAACTTCTTCTCTACTGCTTTCATGACAGAGCGGTTCCGACCATGAGCAGCCGACG 792
 QY 898 CTGAAGGCTTTCTTCAAGTGCAGAGAACCTGTGACAGTCTTACCAATCATTAATCTTT 957
 |||||
 Db 793 CTGAAGGCTTTCTTCAAGTGCAGAGAACCTGTGACAGTCTTACCAATCATTAATCTTT 852
 QY 958 TCCATTAACAAGTAGCCCTGTGATCTGCGCGGAACTCACTGATCAAGATGCTGTGG 1017
 |||||
 Db 853 TCCATTAACAAGTAGCCCTGTGATCTGCGCGGAACTCACTGATCAAGATGCTGTGG 912
 QY 1018 TACAGTATGACAAATAAGTAAAGTATCCCGGTGA 1062
 |||||
 Db 913 TACAGTATGACAAATAAGTAAAGTATCCCGGTGA 957

```

RESULT 15
AL139235 0/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AL139235 Accession AL139235
Fragment Name Begin End
AL139235_0 1 110000
AL139235_1 100001 210000
AL139235_2 200001 310000
AL139235_3 300001 407025
LOCUS AL139235 407025 bp DNA linear HTG 16-JUL-2001
DEFINITION Homo sapiens chromosome 9 clone RP11-3215, 31 unordered pieces.
ACCESSION AL139235
VERSION AL139235.22 GI:11034497
KEYWORDS HTG: HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Plumb,B.
Direct Submission
Submitted (14-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
request: clonerequest@sanger.ac.uk
On Oct 26, 2000 this sequence version replaced gi:11024977.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: ba3215
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 389894 bases at least Q40
Consensus quality: 395933 bases at least Q30
Consensus quality: 400050 bases at least Q20
Insert size: 404025; sum-of-contigs
Insert size: 174695; 17.7% error; agarose-fp
Quality coverage: 4.93x in Q20 bases; sum-of-contigs Quality
coverage: 14.49x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2425: contig of 2425 bp in length
* 2426 2525: gap of 100 bp
* 2526 10023: contig of 7498 bp in length
* 10024 10123: gap of 100 bp
* 10124 36364: contig of 26241 bp in length
* 36365 36464: gap of 100 bp
* 36465 43873: contig of 7409 bp in length
* 43874 43973: gap of 100 bp
* 43974 49988: contig of 6015 bp in length
* 49989 50088: gap of 100 bp
* 50089 93603: contig of 43515 bp in length
* 93604 93703: gap of 100 bp
* 93704 96191: contig of 2488 bp in length
* 96192 96292: gap of 100 bp
* 96292 99433: contig of 3142 bp in length
* 99434 99533: gap of 100 bp
* 99534 101556: contig of 2023 bp in length
* 101557 101656: gap of 100 bp
* 101657 184109: contig of 8253 bp in length
* 184110 184209: gap of 100 bp

```

```

FEATURES
source
*
* 184210 290596: contig of 106487 bp in length
* 290597 290796: gap of 100 bp
* 290797 292892: contig of 2096 bp in length
* 292893 292992: gap of 100 bp
* 292993 296018: contig of 3026 bp in length
* 296019 296118: gap of 100 bp
* 296119 322154: contig of 26036 bp in length
* 322155 322254: gap of 100 bp
* 322255 324646: contig of 2392 bp in length
* 324647 324746: gap of 100 bp
* 324747 329069: contig of 4323 bp in length
* 329070 329169: gap of 100 bp
* 329170 331371: contig of 2202 bp in length
* 331372 331471: gap of 100 bp
* 331472 333909: contig of 2438 bp in length
* 333910 334009: gap of 100 bp
* 334010 369445: contig of 35436 bp in length
* 369446 369545: gap of 100 bp
* 369546 371996: contig of 2451 bp in length
* 371997 372096: gap of 100 bp
* 372097 375925: contig of 3829 bp in length
* 375926 376025: gap of 100 bp
* 376026 378132: contig of 2107 bp in length
* 378133 378232: gap of 100 bp
* 378233 381590: contig of 3358 bp in length
* 381591 381690: gap of 100 bp
* 381691 383726: contig of 2036 bp in length
* 383727 383826: gap of 100 bp
* 383827 390832: contig of 7006 bp in length
* 390833 390932: gap of 100 bp
* 390933 394216: contig of 3284 bp in length
* 394217 394316: gap of 100 bp
* 394317 396358: contig of 2042 bp in length
* 396359 396458: gap of 100 bp
* 396459 398733: contig of 2275 bp in length
* 398734 398833: gap of 100 bp
* 398834 402385: contig of 3552 bp in length
* 402386 402485: gap of 100 bp
* 402486 404522: contig of 2037 bp in length
* 404523 404622: gap of 100 bp
* 404623 407025: contig of 2403 bp in length.
*
* 1..407025
location/Qualifiers
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="9"
  /clone="RP11-3215"
  /clone_1lb="RPC1-11.1"
  1..2425
    /note="assembly fragment:05385
    fragment_chain:1"
  2526..10023
    /note="assembly fragment:03131
    fragment_chain:1"
  10124..36364
    /note="assembly fragment:06402
    fragment_chain:1"
  36465..43873
    /note="assembly fragment:01436
    fragment_chain:1"
  43974..49988
    /note="assembly fragment:04547
    fragment_chain:1"
  50089..93603
    /note="assembly fragment:01636
    fragment_chain:1"
  93704..96191
    /note="assembly fragment:02957
    fragment_chain:2"
  96292..99433
    /note="assembly fragment:00499
    fragment_chain:2"

```

```

misc_feature      99534..101556
                    /note="assembly_fragment:03112
                    fragment_chain:2"
misc_feature      101657..184109
                    /note="assembly_fragment:01419
                    fragment_chain:3"
misc_feature      184210..290696
                    /note="assembly_fragment:00925
                    fragment_chain:3"
misc_feature      290797..292892
                    /note="assembly_fragment:03065
                    fragment_chain:4"
misc_feature      292993..296018
                    /note="assembly_fragment:05057
                    fragment_chain:4"
misc_feature      296119..322154
                    /note="assembly_fragment:00269"
                    fragment_chain:4"
misc_feature      322255..324646
                    /note="assembly_fragment:00686"
                    fragment_chain:4"
misc_feature      324747..329069
                    /note="assembly_fragment:01113"
                    fragment_chain:4"
misc_feature      329170..331371
                    /note="assembly_fragment:01461"
                    fragment_chain:4"
misc_feature      331472..333909
                    /note="assembly_fragment:01516"
                    fragment_chain:4"
misc_feature      334010..369445
                    /note="assembly_fragment:01942.0"
                    fragment_chain:4"
misc_feature      369546..371996
                    /note="assembly_fragment:02194"
                    fragment_chain:4"
misc_feature      372097..375925
                    /note="assembly_fragment:02219"
                    fragment_chain:4"
misc_feature      376026..378132
                    /note="assembly_fragment:02485"
                    fragment_chain:4"
misc_feature      378233..381590
                    /note="assembly_fragment:02649"
                    fragment_chain:4"
misc_feature      381691..383726
                    /note="assembly_fragment:02869"
                    fragment_chain:4"
misc_feature      383827..390832
                    /note="assembly_fragment:03540"
                    fragment_chain:4"
misc_feature      390933..394216
                    /note="assembly_fragment:04405"
                    fragment_chain:4"
misc_feature      394317..396358
                    /note="assembly_fragment:04555"
                    fragment_chain:4"
misc_feature      396459..398733
                    /note="assembly_fragment:05768"
                    fragment_chain:4"
misc_feature      398834..402385
                    /note="assembly_fragment:06106"
                    fragment_chain:4"
misc_feature      402486..404522
                    /note="assembly_fragment:06204"
                    fragment_chain:4"
misc_feature      404623..407025
                    /note="assembly_fragment:06292"
                    fragment_chain:4"
ORIGIN
Query Match      88.3%; Score 938; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 1.1e-216;
Matches 938; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      125 CAGCAATATCTTGACAGTATCATCTCCAGCTGTGSCAAGAGACAGAACTCT 184
DB      35190 CAGCAATATCTTGACAGTATCATCTCCAGCTGTGSCAAGAGACAGAACTCT 35131
QY      185 CCTACAATATCTCTTGGAAGTGTGCTGCGAGTCTTGATCTCTTTTTCATAGTGT 244
DB      35130 CCTACAATATCTCTTGGAAGTGTGCTGCGAGTCTTGATCTCTTTTTCATAGTGT 35071
QY      245 TTGTGACATCTCTGTTGGAAGTATTTGATCTTGAAATGAGATGCTCCAGTCCCGACA 304
DB      35070 TTGTGACATCTCTGTTGGAAGTATTTGATCTTGAAATGAGATGCTCCAGTCCCGACA 35011
QY      305 AGATCATAGAGTGTGATTTCTCATCCATCCACACTCCATATGATTAAGTACCGT 364
DB      35010 AGATCATAGAGTGTGATTTCTCATCCATCCACACTCCATATGATTAAGTACCGT 34951

```

```

QY      365 TAACATTGACAGGTATATGCTGTGCGACCCGCTCAAGTACACAGCGTCTCATACC 424
DB      34950 TAACATTGACAGGTATATGCTGTGCGACCCGCTCAAGTACACAGCGTCTCATACC 34891
QY      425 CAGCCCGACCCGGAAGTCAATTTGATTTTACATCACTCTTCTGACAGCATCC 484
DB      34890 CAGCCCGACCCGGAAGTCAATTTGATTTTACATCACTCTTCTGACAGCATCC 34831
QY      485 CCTATTATCTGTGCGCCCAACATCTGAGCTGGAAGATCAATCAAGACCTGTGACATCAG 544
DB      34830 CCTATTATCTGTGCGCCCAACATCTGAGCTGGAAGATCAATCAAGACCTGTGACATCAG 34771
QY      545 TCTCATCTGATTCACATCTTCAACGCTTACCTGATCTGATCTGATCTTCTTCTCATCT 604
DB      34770 TCTCATCTGATTCACATCTTCAACGCTTACCTGATCTGATCTGATCTTCTTCTCATCT 34711
QY      605 TGAATCATCATTTGTGTAAGAGTCAAGAGAGAGAGCAATTTGCTCTCGTGTACT 664
DB      34710 TGAATCATCATTTGTGTAAGAGTCAAGAGAGAGAGCAATTTGCTCTCGTGTACT 34651
QY      665 CCAAGGGGAAAGACACCGGCATCTGTTTACATTAACCTTCATCTTTGACACATTTGGG 724
DB      34650 CCAAGGGGAAAGACACCGGCATCTGTTTACATTAACCTTCATCTTTGACACATTTGGG 34591
QY      725 CCCCCCGCATCATCATGATTTCTTTACCACTCTATGAGGCGCCCATCCAGAACCGTGGC 784
DB      34590 CCCCCCGCATCATCATGATTTCTTTACCACTCTATGAGGCGCCCATCCAGAACCGTGGC 34531
QY      785 TGTGACATCATGTCGACATTTGCAATGCAATCTAGCCCTTGTGAACAGGCATCACT 844
DB      34530 TGTGACATCATGTCGACATTTGCAATGCAATCTAGCCCTTGTGAACAGGCATCACT 34471
QY      845 TCTTCTCTAGTCTTCAATGCAAGCGGTTCCGACATGCAAGCGCCGACGCTCAAG 904
DB      34470 TCTTCTCTAGTCTTCAATGCAAGCGGTTCCGACATGCAAGCGCCGACGCTCAAG 34411
QY      905 CTTTCTTAAGTGCAGAAAGCAACCTGTACAGTTTCAACCAATCACTTTTCATTA 964
DB      34410 CTTTCTTAAGTGCAGAAAGCAACCTGTACAGTTTCAACCAATCACTTTTCATTA 34351
QY      965 CAAGTACCCCTGATCTGCGCGCAAACTCACTGATCAAGATGCTGTGTACAGT 1024
DB      34350 CAAGTACCCCTGATCTGCGCGCAAACTCACTGATCAAGATGCTGTGTACAGT 34291
QY      1025 ATGACAAAAATGAAAACTTATTAAGTATCCCGTGA 1062
DB      34290 ATGACAAAAATGAAAACTTATTAAGTATCCCGTGA 34253

```

Search completed: December 15, 2004, 00:42:52
 Job time : 3579.97 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:00:04 : Search time 497.373 Seconds
(without alignments)
11208.653 Million cell updates/sec

Title: US-09-995-225B-15

Perfect score: 1062
Sequence: 1 atggagcagacagcagccca.....ctataaagatcccgctga 1062

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: N_Geneseq_23sep04:*
2: geneseqn1980a:*
3: geneseqn1990a:*
4: geneseqn2000a:*
5: geneseqn2001a:*
6: geneseqn2002a:*
7: geneseqn2003a:*
8: geneseqn2004a:*
9: geneseqn2005a:*
10: geneseqn2006a:*
11: geneseqn2007a:*
12: geneseqn2008a:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1062	100.0	1062	6	ABK15563
2	1062	100.0	1062	6	ABK15563
3	1062	100.0	1062	10	ABK15563
4	1060.4	99.8	1062	6	ABK15563
5	1060.4	99.8	1062	6	ABK15563
6	1060.4	99.8	1062	6	ABK15563
7	1060.4	99.8	1062	6	ABK15563
8	1058.8	99.7	1343	10	AA515731
9	1058.8	99.7	1343	10	AA515731
10	1058.8	99.7	1343	12	AD179324
11	1058.8	99.7	1343	12	AD179324
12	1058.8	99.7	1343	12	AD179324
13	1057.4	99.6	1059	6	ABN84273
14	1057.4	99.6	1059	6	ABN84273
15	1050.8	98.9	1526	8	ACA60889
16	1050.8	98.9	1526	8	ACA60889
17	1048.4	98.7	1119	4	AA515730
18	1048.4	98.7	1119	10	AD179322
19	1048.4	98.7	1119	12	AD179322
20	1048.4	98.7	1119	12	AD179322
21	1048.4	98.7	1130	6	AD29672

22	1048.4	98.7	2198	9	ACF05275	ACF05275 Human G-P
23	1039.8	97.9	2189	10	ACF05276	ACF05276 Human G-P
24	1039.8	97.9	2189	10	ADD18120	ADD18120 Human G-P
25	945.2	89.0	1062	6	ABN84273	ABN84273 Human che
26	940.2	88.5	957	12	AD179347	AD179347 NOV10b co
27	940.2	88.5	957	12	AD056026	AD056026 DNA encod
28	936.4	88.2	1002	6	AD27501	AD27501 Human G-P
29	936.4	88.2	1002	6	AD27501	AD27501 Human G-P
30	936.4	88.2	1158	6	AB242880	AB242880 Human GPC
31	936.4	88.2	1032	6	ABN84271	ABN84271 Human che
32	936.4	88.1	1032	6	ABN84270	ABN84270 Human che
33	936.4	88.1	1026	6	ABN84272	ABN84272 Human che
34	934.4	88.0	1000	6	AD28102	AD28102 Human thy
35	917.2	86.4	946	10	AD187879	AD187879 G-coupled
36	844	79.5	1038	8	ABV73373	ABV73373 Mouse TGR
37	844	79.5	1038	12	AD028957	AD028957 Mouse nov
38	842.6	79.3	864	4	AA542811	AA542811 Human G-P
39	828.4	78.0	867	6	AA598134	AA598134 Human DNA
40	828.4	78.0	894	6	AA598143	AA598143 Human DNA
41	809.6	76.2	963	10	AD012695	AD012695 Human GPC
42	758	71.4	795	6	AA598071	AA598071 Human DNA
43	758	71.4	795	8	AB242552	AB242552 Human G-P
44	524	49.3	930	4	AA542808	AA542808 Human G-P
45	268.2	25.3	1023	12	AD028937	AD028937 Mouse nov

ALIGNMENTS

RESULT 1	ABK15563	ABK15563 standard; cDNA; 1062 BP.
ID	ABK15563	
AC	ABK15563	
XX		
DT	08-MAY-2002	(first entry)
XX		
DE	CDNA encoding novel G-protein coupled receptor TGR8, version #2.	
XX		
KW	G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;	
KW	cytotoxic; antiinflammatory; anticancer; fetal brain;	
KW	central nervous system disease; circulatory organ disorder; cancer;	
KW	metabolic disease; immunological disease; gastrointestinal disease;	
XX	gene therapy; transgenic animal; human; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1062
FT		/*tag= a "TGR8"
FT		/product= "G-protein coupled receptor"
FT		/note= "G-protein coupled receptor"
XX		
PN	WO200194582-A1.	
XX		
PD	13-DEC-2001.	
XX		
PF	01-JUN-2001; 2001WO-IP004643.	
XX		
PR	02-JUN-2000; 2000JP-00170446.	
PR	23-JUN-2000; 2000JP-00194926.	
XX		
PA	(TAKE) TAKEDA CHEM IND LTD.	
XX		
PI	Terao Y, Matsui H, Shintani Y,	
XX		
DR	WPI; 2002-164317/21.	
XX		
PT	P-PSDB; AAU76416.	
XX		
PT	Human fetal brain-originated G protein-coupled receptor protein TGR8 and	
PT	encoding DNA, for developing drugs to treat e.g. diseases of the central	
PT	nervous system or circulatory organs, cancer, and metabolic diseases.	
XX		

PS Claim 5; Page 94-95; 102bp; Japanese.
XX
CC The invention describes a human fetal brain-originated G protein-coupled
CC receptor protein, or its salt. The protein and encoded DNA are useful for
CC developing drugs to treat e.g. diseases of the central nervous system or
CC circulatory organs, cancer, metabolic diseases, immunological diseases
CC and gastrointestinal diseases. The invention also describes creation of a
CC probe for gene therapy and construction of a transgenic animal. This
CC sequence encodes the novel G-protein coupled receptor TGR8, described in
CC the method of the invention
XX
SQ Sequence 1062 BP; 240 A; 342 C; 208 G; 272 T; 0 U; 0 Other;
Query Match 100.0%; Score 1062; DB 6; Length 1062;
Best Local Similarity 100.0%; Pred. No. 1.4e-284;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACACAGCAGCCGACCTCGAGCCAAAGCTCGCTGTGTGTGTCCTCCCGGC 60
DB 1 ATGAGACACAGCAGCCGACCTCGAGCCAAAGCTCGCTGTGTGTGTCCTCCCGGC 60
QY 61 TGGGCTGGGGCTTGGGTTTCGTGCGCGGTCTTACTAAGCTCTTGTGCTCGGT 120
DB 61 TGGGCTGGGGCTTGGGTTTCGTGCGCGGTCTTACTAAGCTCTTGTGCTCGGT 120
QY 121 TTACAGCAAAATCTTGAAGATCATCTCTCCAGCTGGTGGCAAGAGACAGAG 180
DB 121 TTACAGCAAAATCTTGAAGATCATCTCTCCAGCTGGTGGCAAGAGACAGAG 180
QY 181 TCTCTCTACAACTATCTTGTGCACTCGTGTGCGGACATCTTGTCTCTTTTCATA 240
DB 181 TCTCTCTACAACTATCTTGTGCACTCGTGTGCGGACATCTTGTCTCTTTTCATA 240
QY 241 GTGTTGTGACTTCCTGTGTGAAGTTTCATCTTGAAGATGCAATGCTTGAAGTCCC 300
DB 241 GTGTTGTGACTTCCTGTGTGAAGTTTCATCTTGAAGATGCAATGCTTGAAGTCCC 300
QY 301 GACAAGATCATGAAGTGTGGAATCTCATCCATCCACACCTCCATATGATTACTGTA 360
DB 301 GACAAGATCATGAAGTGTGGAATCTCATCCATCCACACCTCCATATGATTACTGTA 360
QY 361 CGGTTAACCATTGACAGTATATCGTGTGTGCAACCGCTCAAGTACCAACGCTTCA 420
DB 361 CGGTTAACCATTGACAGTATATCGTGTGTGCAACCGCTCAAGTACCAACGCTTCA 420
QY 421 TACCAAGCCGCGACCCGGAAGTCATGTAAGTGTTAATACCTGCTTCTGACACAG 480
DB 421 TACCAAGCCGCGACCCGGAAGTCATGTAAGTGTTAATACCTGCTTCTGACACAG 480
QY 481 ATCCCTATTACTGTGTGCGCCAAACATCTGAGCTGAAGACTAATCAGACACTGTGAT 540
DB 481 ATCCCTATTACTGTGTGCGCCAAACATCTGAGCTGAAGACTAATCAGACACTGTGAT 540
QY 541 CAGGTCCTATCTGGATCCAGCTGCTTACCGTCTACCTGTGCTCTGCTCTTTTC 600
DB 541 CAGGTCCTATCTGGATCCAGCTGCTTACCGTCTACCTGTGCTCTGCTCTTTTC 600
QY 601 ATCTTGAATCAATATTGTGTGAAGCTCAGAGAGGAAGAATTTTCGTCTCGTGGC 660
DB 601 ATCTTGAATCAATATTGTGTGAAGCTCAGAGAGGAAGAATTTTCGTCTCGTGGC 660
QY 661 TACTCAGGGGAGAGACACCGCATCTTGTTCACCATTAACCTCTTGTGCAACACT 720
DB 661 TACTCAGGGGAGAGACACCGCATCTTGTTCACCATTAACCTCTTGTGCAACACT 720
QY 721 TGGGCCCCCGGATCATGATTTCTTACCACTCTATGAGGGCGCCCATCAGAACCGC 780
DB 721 TGGGCCCCCGGATCATGATTTCTTACCACTCTATGAGGGCGCCCATCAGAACCGC 780
QY 781 TGGGCTGGTGCATCATGTGCGACATGCGACATGCTAGCCCTTTGTAACACAGCATC 840
DB 781 TGGGCTGGTGCATCATGTGCGACATGCGACATGCTAGCCCTTTGTAACACAGCATC 840

QY 841 AACTTCTTCTTACTGCTTCAATGCAAGCGGTTCCGACACATATGACGCGCCACGCTC 900
DB 841 AACTTCTTCTTACTGCTTCAATGCAAGCGGTTCCGACACATATGACGCGCCACGCTC 900
QY 901 AAGGCTTCTTCAATGCGCAGAGAACCTGTACAGTTCTACACCAATCATTAATTTC 960
DB 901 AAGGCTTCTTCAATGCGCAGAGAACCTGTACAGTTCTACACCAATCATTAATTTC 960
QY 961 ATAAAGATGAGCCCTGATCTCGCGGCAAACTACACTGATTAAGATGCTGTGATC 1020
DB 961 ATAAAGATGAGCCCTGATCTCGCGGCAAACTACACTGATTAAGATGCTGTGATC 1020
QY 1021 CAGTATGACAAATAATGAAAACTTAATAAGATCCCGTGA 1062
DB 1021 CAGTATGACAAATAATGAAAACTTAATAAGATCCCGTGA 1062
RESULT 2
ABT04873
ID ABT04873 standard; cDNA, 1062 BP.
XX
AC ABT04873;
XX
DT 11-OCT-2002 (first entry)
XX
DE Human G protein coupled receptor hRUP35 coding sequence.
XX
KW Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
KW hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200242461-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001WO-US044386.
XX
PR 27-NOV-2000; 2000US-0253404P.
PR 12-DEC-2000; 2000US-0253366P.
PR 20-FEB-2001; 2001US-0270266P.
PR 20-FEB-2001; 2001US-0270286P.
PR 06-APR-2001; 2001US-0282032P.
PR 06-APR-2001; 2001US-0282358P.
PR 06-APR-2001; 2001US-0282358P.
PR 06-APR-2001; 2001US-0282365P.
PR 14-MAY-2001; 2001US-0230917P.
PR 31-JUL-2001; 2001US-0309208P.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX
DR WPI; 2002-566565/60.
XX
DR P-PSDB; ABJ04075.
XX
PT Novel endogenous and non-endogenous versions of G protein-coupled
PT receptor useful for identification of candidate compounds as receptor
PT agonists or antagonists for use as therapeutic agents.
XX
PS Claim 31; Page 70; 84bp; English.
XX
CC The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a GPCR coding sequence of the invention
XX
SQ Sequence 1062 BP; 240 A; 342 C; 208 G; 272 T; 0 U; 0 Other;
Query Match 100.0%; Score 1062; DB 6; Length 1062;
Best Local Similarity 100.0%; Pred. No. 1.4e-284;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 1062; DB 10; Length 1062;
 Best Local Similarity 100.0%; Pred. No. 1,4e-284;
 Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAGACACACGACGCCACCTCGAGCCAA CAGTCGCTGTCTTGTGGTCCCGCGGC 60
DB 1 ATGAGACACACGACGCCACCTCGAGCCAA CAGTCGCTGTCTTGTGGTCCCGCGGC 60
QY 61 TCGGCTGCGGCTGGGTTTGGGCCCGGCTCTA CTACAGCTTTGCTGTGCTCGGT 120
DB 61 TCGGCTGCGGCTGGGTTTGGGCCCGGCTCTA CTACAGCTTTGCTGTGCTCGGT 120
QY 121 TTACCGCAAAATATCTTGA CAGTGATCATCTCTCCAGCTGTGGCAAGACAGAG 180
DB 121 TTACCGCAAAATATCTTGA CAGTGATCATCTCTCCAGCTGTGGCAAGACAGAG 180
QY 181 TCTCTCACTATCTTGGCACTCGCTGTGCGA CACTTGTGCTTTTCTATTA 240
DB 181 TCTCTCACTATCTTGGCACTCGCTGTGCGA CACTTGTGCTTTTCTATTA 240
QY 241 GTGTGTGGACCTTCTGTGGAGATTTCATT TTAACATGAGATGCTCAGTCCG 300
DB 241 GTGTGTGGACCTTCTGTGGAGATTTCATT TTAACATGAGATGCTCAGTCCG 300
QY 301 GACAAGATCATAGAAAGCTGGAATTCATCCA TCAACACTTCATATGATTAATGTA 360
DB 301 GACAAGATCATAGAAAGCTGGAATTCATCCA TCAACACTTCATATGATTAATGTA 360
QY 361 CGGTAAACATGACAGGTATATGCTGTGCGA CCGGCTCAAGTACACAGGTCTCA 420
DB 361 CGGTAAACATGACAGGTATATGCTGTGCGA CCGGCTCAAGTACACAGGTCTCA 420
QY 421 TACCCAGCCGCAACCGGAAAGTATGTAAGT TTAACATCACTGCTTCTGACAC 480
DB 421 TACCCAGCCGCAACCGGAAAGTATGTAAGT TTAACATCACTGCTTCTGACAC 480
QY 481 ATCCCTATTAAGTGTGGCCCAACATCTGSA CTGAAGACATCAATGACACTCTGTGCAT 540
DB 481 ATCCCTATTAAGTGTGGCCCAACATCTGSA CTGAAGACATCAATGACACTCTGTGCAT 540
QY 541 CAGCTCTCATCTGATTCACCTGTTACCGTCT ACCTGCGCTCCATCTTCTTC 600
DB 541 CAGCTCTCATCTGATTCACCTGTTACCGTCT ACCTGCGCTCCATCTTCTTC 600
QY 601 ATCTTGAATCAATCATGTGTGTAAGTCAAG TCAAGAGAAAGCAATTTTGTCTCGTGGC 660
DB 601 ATCTTGAATCAATCATGTGTGTAAGTCAAG TCAAGAGAAAGCAATTTTGTCTCGTGGC 660
QY 661 TACTTCCACGGGGAAGACCAACCGCATCTT GTTCAACATTAATCTTTCACACATT 720
DB 661 TACTTCCACGGGGAAGACCAACCGCATCTT GTTCAACATTAATCTTTCACACATT 720
QY 721 TGGGCCCCCGCATCATCATGATCTTTTTCAC ACTCTTATGAGGCGCCATCCAGAACCGC 780
DB 721 TGGGCCCCCGCATCATCATGATCTTTTTCAC ACTCTTATGAGGCGCCATCCAGAACCGC 780
QY 781 TGGCTGTGACATCATGTCCAGATTCGAACAT GTAGCCCTTCTTGAACACAGCCATC 840
DB 781 TGGCTGTGACATCATGTCCAGATTCGAACAT GTAGCCCTTCTTGAACACAGCCATC 840
QY 841 AACTTCTTCTACTGCTTTCATGACGAAGCGGT TCCGACCAATGACGCGCCACGCTC 900
DB 841 AACTTCTTCTACTGCTTTCATGACGAAGCGGT TCCGACCAATGACGCGCCACGCTC 900
QY 901 AAGGCTTTCTTCAATGCGACAGAAACCTGTA CAGTTCTACACCAATCATTAATTTTCC 960
DB 901 AAGGCTTTCTTCAATGCGACAGAAACCTGTA CAGTTCTACACCAATCATTAATTTTCC 960
QY 961 ATAACAATAGCCCTGATCTCGCGGAAATCTCA CTGACATCAAGATGCTGTGTAC 1020
DB 961 ATAACAATAGCCCTGATCTCGCGGAAATCTCA CTGACATCAAGATGCTGTGTAC 1020
QY 1021 CAGTATGACAAATAATGAAACTATATAAGTA TATCCCGTGA 1062
DB 1021 CAGTATGACAAATAATGAAACTATATAAGTA TATCCCGTGA 1062

```

DB 1021 CAGTATGACAAATAATGAAACTATATAAGTATCCCGTGA 1062

```

RESULT 4
ID ABK15562 standard; cDNA; 1062 BP.
XX
AC ABK15562;
XX
DT 08-MAY-2002 (first entry)
XX
DE cDNA encoding novel G-protein coupled receptor TGR8, version #1.
XX
KW G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
KW cytosolic; antiinflammatory; antitumor; fetal brain;
KW central nervous system disease; circulatory organ disorder; cancer;
KW metabolic disease; immunological disease; gastrointestinal disease;
KW gene therapy; transgenic animal; human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1062
FT /tag= a
FT /product= "TGR8"
FT /note= "G-protein coupled receptor"
XX
PN WO200194582-A1.
XX
PD 13-DEC-2001.
XX
PF 01-JUN-2001; 2001WO-JP004643.
XX
PR 02-JUN-2000; 2000JP-00170446.
XX
PR 23-JUN-2000; 2000JP-00194926.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Terao Y, Matsui H, Shintani Y;
XX
DR WPI: 2002-164317/21.
XX
DR P-PSDB; AAU76416.
XX
PT Human fetal brain-originated G-protein-coupled receptor protein TGR8 and
PT encoding DNA, for developing drugs to treat e.g. diseases of the central
PT nervous system or circulatory organs, cancer, and metabolic diseases.
XX
PS Claim 5; Page 94; 102pp; Japanese.
XX
CC The invention describes a human fetal brain-originated G-protein-coupled
CC receptor protein, or its salt. The protein and encoded DNA are useful for
CC developing drugs to treat e.g. diseases of the central nervous system or
CC circulatory organs, cancer, metabolic diseases, immunological diseases
CC and gastrointestinal diseases. The invention also describes creation of a
CC probe for gene therapy and construction of a transgenic animal. This
CC sequence encodes the novel G-protein coupled receptor TGR8, described in
CC the method of the invention.
XX
SQ Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 Other;

```

Query Match 99.8%; Score 1060.4; DB 6; Length 1062;
 Best Local Similarity 99.9%; Pred. No. 3.9e-284;
 Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGACACACGACGCCACCTCGAGCCAA CAGCTGCTGTCTTGTGGTCCCGCGC 60
DB 1 ATGAGACACACGACGCCACCTCGAGCCAA CAGCTGCTGTCTTGTGGTCCCGCGC 60
QY 61 TCGGCTGCGGCTGGGTTTGGGCCCGGCTCTA CTACAGCTTTGCTGTGCTCGGT 120
DB 61 TCGGCTGCGGCTGGGTTTGGGCCCGGCTCTA CTACAGCTTTGCTGTGCTCGGT 120

```


QY 121 TTACAGCAAAATATCTTGACAGTATCATCTCCAGCTGGGAGAGACAGAG 180
 DB 121 TTACAGCAAAATATCTTGACAGTATCATCTCCAGCTGGGAGAGACAGAG 180
 QY 181 TCCTCCACAAATATCTTGACAGTATCATCTCCAGCTGGGAGAGACAGAG 240
 DB 181 TCCTCCACAAATATCTTGACAGTATCATCTCCAGCTGGGAGAGACAGAG 240
 QY 241 GTGTGTGTGAGATCTCTGTGTGAGAGATTTATCTTGAACTGAGATGCTCAGGTCCC 300
 DB 241 GTGTGTGTGAGATCTCTGTGTGAGAGATTTATCTTGAACTGAGATGCTCAGGTCCC 300
 QY 301 GACAGATCATAGAGTGTGAGATTTCTCATTCATCCACACTCCATAGATTTACTGTA 360
 DB 301 GACAGATCATAGAGTGTGAGATTTCTCATTCATCCACACTCCATAGATTTACTGTA 360
 QY 361 CCGTTAACATGACAGATATATCGCTGTGACACCGCTCAAGTACACACGGTCTCA 420
 DB 361 CCGTTAACATGACAGATATATCGCTGTGACACCGCTCAAGTACACACGGTCTCA 420
 QY 421 TACCCAGCCCGCACCCGGAAGTCAATTTGTAAGTGTTAATCACTGCTCTGACAGC 480
 DB 421 TACCCAGCCCGCACCCGGAAGTCAATTTGTAAGTGTTAATCACTGCTCTGACAGC 480
 QY 481 ATCCCATATTAATGAGTGTGAGATTTCTCATTCATCCACACTCCATAGATTTACTGTA 540
 DB 481 ATCCCATATTAATGAGTGTGAGATTTCTCATTCATCCACACTCCATAGATTTACTGTA 540
 QY 541 CAGCTCTCATCTGATATCCACTGCTTCAACGCTTACCTGAGTGTGCTCATCTTCTTC 600
 DB 541 CAGCTCTCATCTGATATCCACTGCTTCAACGCTTACCTGAGTGTGCTCATCTTCTTC 600
 QY 601 ATTGTGAATCATCAATTTGTGTAACAGTCTGAGAGAGAGAGCAATTTTCTGCTGGC 660
 DB 601 ATTGTGAATCATCAATTTGTGTAACAGTCTGAGAGAGAGAGCAATTTTCTGCTGGC 660
 QY 661 TACTCAGCGGGAGAGACACCGGATCTGTGACCAATTAACCTCAATCTTGTGACAGATT 720
 DB 661 TACTCAGCGGGAGAGACACCGGATCTGTGACCAATTAACCTCAATCTTGTGACAGATT 720
 QY 721 TGGGCCCCCGGATCATCATGATTTCTTACCACTGATGGGGCGCCATCCAGAACCGC 780
 DB 721 TGGGCCCCCGGATCATCATGATTTCTTACCACTGATGGGGCGCCATCCAGAACCGC 780
 QY 781 TGGCTGTGACATCATGTCGACATTTGCCACATGCTGACCTTCTGAAACAGCCATC 840
 DB 781 TGGCTGTGACATCATGTCGACATTTGCCACATGCTGACCTTCTGAAACAGCCATC 840
 QY 841 AACTTCTCTCTACAGCTTCAATCAGCAAGCGGTTCCGACATGGCAGCCGACGCTC 900
 DB 841 AACTTCTCTCTACAGCTTCAATCAGCAAGCGGTTCCGACATGGCAGCCGACGCTC 900
 QY 901 AAGGCTTTCTTCAAGTGTGAGAGAGCAACCTGTACAGTTCATCAACATCAATCTTTCC 960
 DB 901 AAGGCTTTCTTCAAGTGTGAGAGAGCAACCTGTACAGTTCATCAACATCAATCTTTCC 960
 QY 961 ATPAACAAGTATGAGCCCTGAGATCTGCGCGGCAACTCAACTGCAATCAAGATGTGTATC 1020
 DB 961 ATPAACAAGTATGAGCCCTGAGATCTGCGCGGCAACTCAACTGCAATCAAGATGTGTATC 1020
 QY 1021 CAGTATGACAAATATGAGAAACCTATTAAGTATCCCGGTA 1062
 DB 1021 CAGTATGACAAATATGAGAAACCTATTAAGTATCCCGGTA 1062
 RESULT 5
 ACC71785
 XX ACC71785 standard; DNA; 1062 BP.
 AC ACC71785;
 XX
 XX 25-JUL-2003 (first entry)
 XX

DE Human G protein-coupled receptor DNA #SEQ ID 1.
 XX Human; G protein-coupled receptor; anorectic; eating disorder; obesity;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..1062
 FT CDS /tag= a
 FT /product= "G protein-coupled receptor"
 XX
 PN W02003027142-A1.
 XX
 PD 03-APR-2003.
 XX
 PF 19-SEP-2002; 2002W0-JP009626.
 XX
 PR 21-SEP-2001; 2001JP-00286278.
 XX
 PA (YAMA) YAMANOUCHI PHARM CO LTD.
 XX
 PI Matsumoto S, Takasaki J, Kurama T, Saito T, Kamohara M, Soga T;
 PI HiYama H;
 XX
 DR WPI; 2003-333291/31.
 XX
 DR P-PSDB; ABR4457.
 XX
 PS New G protein-coupled receptor and encoded gene, useful in screening
 PT preventives or remedies for eating disorders or obesity.
 PT
 PS Disclosure; Page 31-33; 48bp; Japanese.
 CC The invention relates to a novel G protein-coupled receptor. The protein
 CC and its encoded gene are useful for screening preventives or remedies for
 CC eating disorders or obesity. The current sequence represents a G protein
 CC coupled receptor encoding sequence
 CC
 XX
 SQ Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 Other;
 Query Match 99.8%; Score 1060.4; DB 10; Length 1062;
 Best Local Similarity 99.9%; Pred. No. 3.9e-284;
 Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGAGACACACGACGCCACCTCCGACGCAACAGCTGCTGTGTGTGTCCTCCGCGC 60
 DB 1 ATGAGACACACGACGCCACCTCCGACGCAACAGCTGCTGTGTGTGTCCTCCGCGC 60
 QY 61 TGGGCTGTGGGCTTGGGTTTGGTGGCCGCTGCTACTAGGCTCTTGTGCTGCTCGGT 120
 DB 61 TGGGCTGTGGGCTTGGGTTTGGTGGCCGCTGCTACTAGGCTCTTGTGCTGCTCGGT 120
 QY 121 TTACAGCAAAATATCTTGACAGTATCATCTCCAGCTGGGAGAGACAGAG 180
 DB 121 TTACAGCAAAATATCTTGACAGTATCATCTCCAGCTGGGAGAGACAGAG 180
 QY 181 TCCTCCACAAATATCTTGACAGTATCATCTCCAGCTGGGAGAGACAGAG 240
 DB 181 TCCTCCACAAATATCTTGACAGTATCATCTCCAGCTGGGAGAGACAGAG 240
 QY 241 GTGTGTGTGAGATCTCTGTGTGAGAGATTTATCTTGAACTGAGATGCTCAGGTCCC 300
 DB 241 GTGTGTGTGAGATCTCTGTGTGAGAGATTTATCTTGAACTGAGATGCTCAGGTCCC 300
 QY 301 GACAGATCATAGAGTGTGAGATTTCTCATTCATCCACACTCCATAGATTTACTGTA 360
 DB 301 GACAGATCATAGAGTGTGAGATTTCTCATTCATCCACACTCCATAGATTTACTGTA 360
 QY 361 CCGTTAACATGACAGATATATCGCTGTGACACCGCTCAAGTACACACGGTCTCA 420
 DB 361 CCGTTAACATGACAGATATATCGCTGTGACACCGCTCAAGTACACACGGTCTCA 420
 QY 421 TACCCAGCCCGCACCCGGAAGTCAATTTGTAAGTGTTAATCACTGCTCTGACAGC 480

```
Db      |||||
421  TACCCAGCCGCGACCGGAAAGTCATGTAGGTTTACATCACTGCTCTGACCAAGC 480
Qy      |||||
481  ATCCCTATTATCTGGTGGCCCAACATCTGCACTGAAAGACTATACAGACCTCTGGAT 540
Db      |||||
481  ATCCCTATTATCTGGTGGCCCAACATCTGCACTGAAAGACTATACAGACCTCTGGAT 540
Qy      |||||
541  CAGCTCTCATCTGATTCATCTGCTTCAACCGTCTACCTGGTCCCTGCTCATCTTCTTC 600
Db      |||||
541  CAGCTCTCATCTGATTCATCTGCTTCAACCGTCTACCTGGTCCCTGCTCATCTTCTTC 600
Qy      |||||
601  ATCTTGAACCTCAATCAATGTGTGTAACAAGCTCAGAGGAGAGCAATTTGTCTCGTGGC 660
Db      |||||
601  ATCTTGAACCTCAATCAATGTGTGTAACAAGCTCAGAGGAGAGCAATTTGTCTCGTGGC 660
Qy      |||||
661  TACTCAGCGGGAAGACACCGGCATCTGTTACCATTAACCTCCATCTTTCGACACTT 720
Db      |||||
661  TACTCAGCGGGAAGACACCGGCATCTGTTACCATTAACCTCCATCTTTCGACACTT 720
Qy      |||||
721  TGGGCCCCCGGCATCATCATGATTTCTTACCACTTATGGGGCGCCCATCCAGAACGGC 780
Db      |||||
721  TGGGCCCCCGGCATCATCATGATTTCTTACCACTTATGGGGCGCCCATCCAGAACGGC 780
Qy      |||||
781  TGGCTGTGTGACATCATATGTCCGACATTTGCCAAGATGCTGACCTTCTGAAACAGCCATC 840
Db      |||||
781  TGGCTGTGTGACATCATATGTCCGACATTTGCCAAGATGCTGACCTTCTGAAACAGCCATC 840
Qy      |||||
841  AACCTTCTCTCTACTGCTTCTCATCAGCAAGCGGTTCGCAACATGAGCGCGCAAGCGCTC 900
Db      |||||
841  AACCTTCTCTCTACTGCTTCTCATCAGCAAGCGGTTCGCAACATGAGCGCGCAAGCGCTC 900
Qy      |||||
901  AAGGCTTCTTCTCAAGTCCAGAAAGCACTGTGACAGTTTACCAATCAATCAATCTTTTCC 960
Db      |||||
901  AAGGCTTCTTCTCAAGTCCAGAAAGCACTGTGACAGTTTACCAATCAATCAATCTTTTCC 960
Qy      |||||
961  ATTAACAGTGAAGCCCTGGATCTGCGCGGCAAACTCACTGTCATCAAGATGCTGTGTATC 1020
Db      |||||
961  ATTAACAGTGAAGCCCTGGATCTGCGCGGCAAACTCACTGTCATCAAGATGCTGTGTATC 1020
Qy      |||||
1021  CAGTATGACAAATAATGAAAACTTATATAAGTATCCCGGTGA 1062
Db      |||||
1021  CAGTATGACAAATAATGAAAACTTATATAAGTATCCCGGTGA 1062

RESULT 6
ABV73364
ID      ABV73364 standard; DNA; 1202 BP.
XX      ABV73364;
AC      ABV73364;
XX      22-JAN-2003 (first entry)
DT      Human TGR20 polypeptide encoding DNA.
DE      Human TGR20 polypeptide encoding DNA.
XX      G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory;
KW      antilipemic; neuroprotective; antidiabetic; antidiabetic; human;
KW      antiparkinsonian; antilipemic; TGR20; gene; de.
XX      Homo sapiens.
OS      Homo sapiens.
XX      Key
FH      Location/Qualifiers
FT      68..1129
FT      CDS
FT      /tag=
FT      /product= "TGR20"
FT      /note= "GPCR polypeptide"
PN      MO200277001-A2.
XX      03-OCT-2002.
PD      08-MAR-2002; 2002MO-US007171.
XX      09-MAR-2001; 2001US-00802803.
PR
```

```
PR      16-MAR-2001; 2001US-0276649P.
XX      (TULAR) TULARIK INC.
XX      Tian H, Zhao J, Chen J, Cutler G;
XX      WPI; 2003-018881/01.
XX      P-PSDB; ABB82499.
XX      New G-protein coupled receptor polypeptides and polynucleotides useful
XX      for identifying compounds for treating a TGR-associated disorder, e.g.
XX      psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's
XX      disease, anemia.
XX      Claim 6, Page 63; 87pp; English.
XX      The invention relates to G-protein coupled receptor (GPCR) polypeptides
XX      and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183,
XX      TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are
XX      useful for identifying compounds for treating a TGR-associated disorder,
XX      such as psoriasis, inflammatory bowel disease, hyperlipidemia,
XX      Parkinson's disease, Huntington's disease, anemia, immune and blood
XX      disorders, ulcerative colitis, Crohn's disease or spleen enlargement.
XX      They are also useful for identifying cells such as kidney, liver,
XX      hypothalamus, colon, adipose, or spleen cells, for forensic and
XX      paternity determination, diagnosing diseases and examining signal
XX      transduction. The present sequence represents a human TGR20 polypeptide
XX      encoding DNA
SQ      Sequence 1202 BP; 267 A; 380 C; 258 G; 297 T; 0 U; 0 Other;
XX      Query Match
XX      Best Local Similarity 99.8%; Score 1060.4; DB 8; Length 1202;
XX      Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1 ATGAGCAGCAGCCGACCCGACCTGCGACGCAACAGCTGGCTGTGGTGGTCCCGGCG 60
Db      68 ATGAGCAGCAGCCGACCCGACCCGACCTGCGACGCAACAGCTGGCTGTGGTGGTCCCGGCG 127
Qy      61 TGGGCTGGGCTGGGTTGGTGGTGGCGGTCATCAGACCTTCTGCTGCTGGT 120
Db      128 TGGGCTGGGCTGGGTTGGTGGTGGCGGTCATCAGACCTTCTGCTGCTGGT 187
Qy      121 TTACGAGCAATATCTTGAACAGTATCTCTTCCAGCTGTGTGGCAAGAGACAGAG 180
Db      188 TTACGAGCAATATCTTGAACAGTATCTCTTCCAGCTGTGTGGCAAGAGACAGAG 247
Qy      181 TCTCTTACACTATCTCTTGGCACTGCTGCTGCGCAACCTTGTGCTCTTTTCTATA 240
Db      248 TCTCTTACACTATCTCTTGGCACTGCTGCTGCGCAACCTTGTGCTCTTTTCTATA 307
Qy      241 GTGTTGTGAGCTCTGTTGGAGATTCTTGAACATGCAAGATGCTTCAAGTCCG 300
Db      308 GTGTTGTGAGCTCTGTTGGAGATTCTTGAACATGCAAGATGCTTCAAGTCCG 367
Qy      301 GACAGATCATGAGTGTGGAATTCTCATATCCACACTCCATATGATATGATATCTGTA 360
Db      368 GACAGATCATGAGTGTGGAATTCTCATATCCACACTCCATATGATATGATATCTGTA 427
Qy      361 CGCTTAACATGACAGGATATGCTGTCTGCAACCGGCTCAAGTACCAAGGCTTCA 420
Db      428 CGCTTAACATGACAGGATATGCTGTCTGCAACCGGCTCAAGTACCAAGGCTTCA 487
Qy      421 TACCCAGCCGCGACCGGAAAGTCATGTAGGTTTACATCACTGCTTCTGACCAAGC 480
Db      488 TACCCAGCCGCGACCGGAAAGTCATGTAGGTTTACATCACTGCTTCTGACCAAGC 547
Qy      481 ATCCCTATTATCTGGTGGCCCAACATCTGCACTGAAAGACTATACAGACCTCTGGAT 540
Db      548 ATCCCTATTATCTGGTGGCCCAACATCTGCACTGAAAGACTATACAGACCTCTGGAT 607
Qy      541 CAGTCTCATCTGATTCATCTGCTTCAACCGTCTACCTGGTCCCTGCTCATCTTCTTC 600
```

Db 608 CAGCTCTCATCTGATCACTGCTTACCGCTCTACCTGGGCGCTCCATCTTCTTC 667
 Qy 601 ATCTTGAACTCATCTATTGTGTACAGCTCAGAGGAGGCAATTTTCTCTCCGTGC 660
 Db 668 ATCTTGAACTCATCTATTGTGTACAGCTCAGAGGAGGCAATTTTCTCTCCGTGC 727
 Qy 661 TACTCCAGGGGAGAGCAGCGCATCTTTGTCACATTCCTCCATCTTTGCGACACT 720
 Db 728 TACTCCAGGGGAGAGCAGCGCATCTTTGTCACATTCCTCCATCTTTGCGACACT 787
 Qy 721 TGGGCCCCCGCATCATCTGATTTCTTTACCACTCTATGGGGGCCCATCAGAACCG 780
 Db 788 TGGGCCCCCGCATCATCTGATTTCTTTACCACTCTATGGGGGCCCATCAGAACCG 847
 Qy 781 TGGCTGGTGACATCATCTGTCGACATTTGGCAATGTTAGCCCTTCTGAAACAAGCATC 840
 Db 848 TGGCTGGTGACATCATCTGTCGACATTTGGCAATGTTAGCCCTTCTGAAACAAGCATC 907
 Qy 841 AACTCTCTCTACTGCTTCATCAGCAGCGGTTCCGACATGGAGCGCGCAGGCTC 900
 Db 908 AACTCTCTCTACTGCTTCATCAGCAGCGGTTCCGACATGGAGCGCGCAGGCTC 967
 Qy 901 AAGCTTTCTTCAAGTCCAGAGCAACTCTGTACAGTTCTACCACTCATTAATTTTCC 960
 Db 968 AAGCTTTCTTCAAGTCCAGAGCAACTCTGTACAGTTCTACCACTCATTAATTTTCC 1027
 Qy 961 ATACCAAGTACGCTCTGATCTGCGCGCAACTCACACTGCATCAAGATGCTGTGTAC 1020
 Db 1028 ATACCAAGTACGCTCTGATCTGCGCGCAACTCACACTGCATCAAGATGCTGTGTAC 1087
 Qy 1021 CAGTATGACAAAATGGAATACTATTAAGTATCCCGCTGA 1062
 Db 1088 CAGTATGACAAAATGGAATACTATTAAGTATCCCGCTGA 1129
 RESULT 7
 ADF70587
 ID ADF70587 standard; DNA; 1776 BP.
 AC ADF70587;
 DT 12-FEB-2004 (first entry)
 XX Orphan receptor ligand-related human protein gene SeqID210.
 DE Orphan receptor ligand-related human protein gene SeqID210.
 XX
 KW ligand; orphan receptor protein; fusion protein; fluorescent protein;
 KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
 KW GFPuv; Enhanced GFP; EGFP; human; gene; de.
 OS Homo sapiens.
 XX
 PN WO2003071272-A1.
 XX
 PD 28-AUG-2003.
 XX
 PF 21-FEB-2003; 2003WO-JP001901.
 XX
 PR 22-FEB-2002; 2002JP-00045728.
 PR 23-JUL-2002; 2002JP-00213949.
 PR 11-OCT-2002; 2002JP-00298237.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M,
 XX
 DR MPI; 2003-697654/56.
 DR P-PDB; ADF70485.
 XX
 PT Transformation of cells with a fusion protein of an orphan receptor
 PT protein with a fluorescent protein useful for identification of ligands
 PT to the orphan receptor.
 XX
 PS Example 4; SEQ ID NO 210; 594bp; Japanese.

XX This invention relates to a novel method of identifying ligands to an
 CC orphan receptor protein which comprises transforming cells with DNA
 CC encoding a fusion protein of the orphan receptor with a fluorescent
 CC protein, so that the fusion protein is expressed in the cells (or cell
 CC membranes isolated from them) and contacting the cells with the potential
 CC ligand to be tested. A suitable fluorescent protein for incorporation in
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
 CC identification of ligands binding to an orphan receptor protein.
 XX
 SQ Sequence 1776 BP; 471 A; 500 C; 358 G; 447 T; 0 U; 0 Other;
 Query Match 99.8%; Score 1060; DB 10; Length 1776;
 Best Local Similarity 100.0%; Pred. No. 6, 4e-284;
 Matches 1060; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGAGACACACGACGCCCACTCGACAGCAAGCTCGCTGTGTGTCCTCCCGGC 60
 Db 1 ATGAGACACACGACGCCCACTCGACAGCAAGCTCGCTGTGTGTCCTCCCGGC 60
 Qy 61 TCGGCTGGGCTTGGGTTTGGTCCGCTGCTACTACAGCTCTTGCTGTGCTCGGT 120
 Db 61 TCGGCTGGGCTTGGGTTTGGTCCGCTGCTACTACAGCTCTTGCTGTGCTCGGT 120
 Qy 121 TTACGACAAATTTCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAACAGAA 180
 Db 121 TTACGACAAATTTCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAACAGAA 180
 Qy 121 TTACGACAAATTTCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAACAGAA 180
 Db 121 TTACGACAAATTTCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAACAGAA 180
 Qy 181 TCCCTCTACAACTATCTTTGGCACTCGCTGCTGCGAGACTTGTGCTCTTTTCAT 240
 Db 181 TCCCTCTACAACTATCTTTGGCACTCGCTGCTGCGAGACTTGTGCTCTTTTCAT 240
 Qy 181 TCCCTCTACAACTATCTTTGGCACTCGCTGCTGCGAGACTTGTGCTCTTTTCAT 240
 Db 181 TCCCTCTACAACTATCTTTGGCACTCGCTGCTGCGAGACTTGTGCTCTTTTCAT 240
 Qy 241 GTGTTTGTGACTTCTGTGTGGAAGATTTCTTGAACATGAGATGCTTCAGTCCC 300
 Db 241 GTGTTTGTGACTTCTGTGTGGAAGATTTCTTGAACATGAGATGCTTCAGTCCC 300
 Qy 241 GTGTTTGTGACTTCTGTGTGGAAGATTTCTTGAACATGAGATGCTTCAGTCCC 300
 Db 241 GTGTTTGTGACTTCTGTGTGGAAGATTTCTTGAACATGAGATGCTTCAGTCCC 300
 Qy 301 GACAAATCATGAAGTGTGGAATTCATCATCATCACAACCTCCATATGATTAATCT 360
 Db 301 GACAAATCATGAAGTGTGGAATTCATCATCATCACAACCTCCATATGATTAATCT 360
 Qy 301 GACAAATCATGAAGTGTGGAATTCATCATCATCACAACCTCCATATGATTAATCT 360
 Db 301 GACAAATCATGAAGTGTGGAATTCATCATCATCACAACCTCCATATGATTAATCT 360
 Qy 361 CGGTAAACATGACAGATATCGCTGTGCGACCCGCTCAAGTACACAGGCTCTCA 420
 Db 361 CGGTAAACATGACAGATATCGCTGTGCGACCCGCTCAAGTACACAGGCTCTCA 420
 Qy 421 TACCCAGCCCGACCCGGAAGTCAATGTGAAGTTTACATCACTGCTTCTGACGAG 480
 Db 421 TACCCAGCCCGACCCGGAAGTCAATGTGAAGTTTACATCACTGCTTCTGACGAG 480
 Qy 481 ATCCCTATTTACTGTGTGCGCCCAACATCTGACATGAAGATCACTACCTCTGTGCA 540
 Db 481 ATCCCTATTTACTGTGTGCGCCCAACATCTGACATGAAGATCACTACCTCTGTGCA 540
 Qy 481 ATCCCTATTTACTGTGTGCGCCCAACATCTGACATGAAGATCACTACCTCTGTGCA 540
 Db 481 ATCCCTATTTACTGTGTGCGCCCAACATCTGACATGAAGATCACTACCTCTGTGCA 540
 Qy 541 CAGTCTCATCTGATTCACAGCTTCACTGCTTACCTGCTGCTGCTCATCTTCTTC 600
 Db 541 CAGTCTCATCTGATTCACAGCTTCACTGCTTACCTGCTGCTGCTCATCTTCTTC 600
 Qy 601 ATCTTGAACTCATCTATTGTGTACAGCTCAGAGGAGAGCAATTTTCTCCGTGCG 660
 Db 601 ATCTTGAACTCATCTATTGTGTACAGCTCAGAGGAGAGCAATTTTCTCCGTGCG 660
 Qy 601 ATCTTGAACTCATCTATTGTGTACAGCTCAGAGGAGAGCAATTTTCTCCGTGCG 660
 Db 601 ATCTTGAACTCATCTATTGTGTACAGCTCAGAGGAGAGCAATTTTCTCCGTGCG 660
 Qy 661 TACTCCAGGGGAGAGCAGCGCATCTTTGTCACATTCCTCCATCTTTGCGACACT 720
 Db 661 TACTCCAGGGGAGAGCAGCGCATCTTTGTCACATTCCTCCATCTTTGCGACACT 720
 Qy 721 TGGGCCCCCGCATCATCTGATTTCTTTACCACTCTATGGGGGCCCATCAGAACCG 780
 Db 721 TGGGCCCCCGCATCATCTGATTTCTTTACCACTCTATGGGGGCCCATCAGAACCG 780
 Qy 781 TGGCTGGTGACATCATCTGTCGACATTTGGCAATGTTAGCCCTTCTGAAACAAGCATC 840
 Db 781 TGGCTGGTGACATCATCTGTCGACATTTGGCAATGTTAGCCCTTCTGAAACAAGCATC 840

QY 841 AACTTCTTCTCTACTGCTTTCATCAGCAGAGGGTTCGACCAATGGACGGCCGACGCTC 900
 DB 841 AACTTCTTCTCTACTGCTTTCATCAGCAGAGGGTTCGACCAATGGACGGCCGACGCTC 900
 QY 901 AAGGCTTCTTCAAGTCCGAGAGCAACCTGTACAGTTCTTACCAATCAATCTTTCC 960
 DB 901 AAGGCTTCTTCAAGTCCGAGAGCAACCTGTACAGTTCTTACCAATCAATCTTTCC 960
 QY 961 ATAACAAGTAGCCCTGGATCTGGCGGCAACCTGCACATGCATCAAGATGCTGTAC 1020
 DB 961 ATAACAAGTAGCCCTGGATCTGGCGGCAACCTGCACATGCATCAAGATGCTGTAC 1020
 QY 1021 CAGTATGACAAATATGAAATCTATTAAGTATCCCGT 1060
 DB 1021 CAGTATGACAAATATGAAATCTATTAAGTATCCCGT 1060
 RESULT 8
 ID AAS15731 standard; cDNA; 1343 BP.
 AC AAS15731;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE DNA encoding chemokine receptor family related protein, NOV11.
 XX
 KM NOV; cytostatic; psoriasis; neurotropic; neuroprotectant;
 KM cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;
 KM haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;
 KM differentiation; proliferation; haematopoiesis; wound healing;
 KM angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;
 KM haemophilia; allergy; Pendred syndrome; skeletal dysplasia;
 KM ischemic injury; neuroepithelial disorder; hepatitis; heart failure;
 KM chemokine receptor; chromosome 1; 8s.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1 /*tag= a
 FT 5'UTR 2.1063
 FT CDS /*tag= b
 FT 2.142 /product= "Chemokine receptor related protein, NOV11"
 FT sig_peptide /*tag= c
 FT 143..1060
 FT mat_peptide /*tag= d
 FT /note= "Mature chemokine receptor related protein, NOV11"
 FT 3'UTR /*tag= e
 FT 1064..1343
 FT
 PN WO200170978-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 20-MAR-2001; 2001WO-US009093.
 XX
 PR 20-MAR-2000; 2000US-019076BP.
 PR 20-MAR-2000; 2000US-019083BP.
 PR 22-MAR-2000; 2000US-0190972P.
 PR 22-MAR-2000; 2000US-0191199P.
 PR 24-MAR-2000; 2000US-0191947P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 29-MAR-2000; 2000US-0192836P.
 PR 31-MAR-2000; 2000US-0193843P.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX Taupier RJ, Majumder K, Spaderma SK, Smithson G, Mezeas PS;

PI Vernet CAM;
 XX
 DR WPI; 2001-639127/73.
 DR P-PSDB; NAUI0068.
 XX
 PT Polypeptides and nucleic acids related to chloride channel, insulin-like
 PT growth factor family of proteins, useful for diagnosing and treating
 PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
 PS
 PS Claim 9; Page 44; 151pp; English.
 XX
 CC The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX
 CC polypeptides are useful for treating pathology associated with NOVX
 CC polypeptide, determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX, identifying agents binding to
 CC NOVX and treatment of disorders associated with altered expression of
 CC members of chloride channel-associated proteins e.g. cystic fibrosis and
 CC congenital myotonia. NOVX proteins are useful in treatment of disorders
 CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,
 CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
 CC in one or more organs (e.g. haemophilia, anaemia), Pendred syndrome,
 CC skeletal dysplasias, disorders characterised by altered cell shape,
 CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial
 CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
 CC treatment of disorders of vascular smooth muscle cell differentiation,
 CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are
 CC useful to screen for molecules which inhibit or enhance NOVX activity or
 CC function and are useful as targets for the identifying small molecules,
 CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
 CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV
 CC sequences are also useful for: identifying a cell or tissue type in a
 CC biological sample; amplifying DNA sequences from very small biological
 CC samples e.g. hair or skin or body fluids and as primers and probes to
 CC identify and/or clone NOVX homologues. NOVX proteins are useful
 CC immunogens to generate antibodies to monitor protein levels and modulate
 CC NOVX activity. Cells comprising the nucleic acids are useful for
 CC producing transgenic animals, for studying the function and/or activity
 CC of NOVX protein and identifying and/or evaluating modulators of NOVX
 CC protein activity. This sequence encodes NOV11 (located on chromosome 1)
 CC related to the chemokine receptor family of proteins, one of 12 NOV
 CC polypeptides described in the method of the invention
 XX
 SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
 Query Match 99.7%; Score 1058.8; DB 4; Length 1343;
 Beet Local Similarity 99.8%; Pred. No. 1.2e-283;
 Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGGAGCAGACGACGCCACCTCGACGCCAAGCTCGTGTGTGTGTCCTCCCGGC 60
 DB 2 ATGGAGCAGACGACGCCACCTCGACGCCAAGCTCGTGTGTGTGTCCTCCCGGC 61
 QY 61 TCGGCTTCGGGCTTGGTTCGTCGGTGTCTTACTACAGCCTTGTGCTGCTGGT 120
 DB 62 TCGGCTTCGGGCTTGGTTCGTCGGTGTCTTACTACAGCCTTGTGCTGCTGGT 121
 QY 121 TTACGAGCAATATCTTGACAGTATCATCTCTCCAGCTGGGGAAGAAGCAGAAG 180
 DB 122 TTACGAGCAATATCTTGACAGTATCATCTCTCCAGCTGGGGAAGAAGCAGAAG 181
 QY 181 TCTCTCTACCAATATCTTGGCACTGCTGTCGCGACATCTTGTGCTCTTTTTCATA 240
 DB 182 TCTCTCTACCAATATCTTGGCACTGCTGTCGCGACATCTTGTGCTCTTTTTCATA 241
 QY 241 GTGTTTGTGACTTCCCTGTGGAGATTTTATCTTGAACATGACAGATGCTCAGGTCCC 300
 DB 242 GTGTTTGTGACTTCCCTGTGGAGATTTTATCTTGAACATGACAGATGCTCAGGTCCC 301
 QY 301 GACAAGATCATAGAGAGTGGAAATTTCTCATCCATCCAGACCTTCATATGATTAATCTGTA 360
 DB 302 GACAAGATCATAGAGAGTGGAAATTTCTCATCCATCCAGACCTTCATATGATTAATCTGTA 361
 QY 361 CCGTTAACATTTGACAGATATATGCTGTCTGCGACCCGCTCAAGTACACACGCTCTCA 420

```

Db      362  CCCTTAACATTAAGATATATACAGTCTGTCCACCCCGTCAAGTACCAAGGCTTCA 421
Qy      421  TACCCAGCCCGCAACCCGAAAGTCATTGTAAGTGTATACATCACTGCTTCTGACCAAGC 480
Db      422  TACCCAGCCCGCAACCCGAAAGTCATTGTAAGTGTATACATCACTGCTTCTGACCAAGC 481
Qy      481  ATCCCTTATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db      482  ATCCCTTATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
Qy      541  CAGCTCTCATCTGATATCCACTGCTTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db      542  CAGCTCTCATCTGATATCCACTGCTTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
Qy      601  ATCTTGAATCAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db      602  ATCTTGAATCAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661
Qy      661  TACTTCAACGGGGAAGACCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db      662  TACTTCAACGGGGAAGACCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
Qy      721  TGGGCCCCCGCATCATCATGATATCTTTTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      722  TGGGCCCCCGCATCATCATGATATCTTTTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781
Qy      781  TGGCTGTGTCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db      782  TGGCTGTGTCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 841
Qy      841  AACTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 900
Db      842  AACTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 901
Qy      901  AAGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960
Db      902  AAGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 961
Qy      961  ATTAACAGTAGAGCCCTGATCTGCGCGCAAACTCACTGATCAAGATGCTGATGCTGATC 1020
Db      962  ATTAACAGTAGAGCCCTGATCTGCGCGCAAACTCACTGATCAAGATGCTGATGCTGATC 1021
Qy      1021  CAGTATGACAAAATGAAAACTATTAAGTATCCCGCTGA 1062
Db      1022  CAGTATGACAAAATGAAAACTATTAAGTATCCCGCTGA 1063

```

RESULT 9
ADJ87768 standard, DNA: 1343 BP.

```

XX      06-MAY-2004 (first entry)
XX      G-coupled protein receptor-related protein coding sequence #57.
XX      novel protein; G-coupled protein receptor-related protein;
XX      cardiomypathy; atherosclerosis; cell signal processing-related disorder;
XX      metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
XX      Huntington's disease; epilepsy; anxiety; pain; hypercholesterolaemia;
XX      obesity; hyperemion; Crohn's disease; systemic lupus erythematosus;
XX      viral infections; bacterial infection; parasitic infection;
XX      hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;
XX      Alzheimer's disease; tuberosus sclerosis; hypercalcaemia; cerebral palsy;
XX      gene; de.
XX      Unidentified.
XX      WO2002102321-A2.
XX      27-DEC-2002.

```

```

XX      18-JUN-2002; 2002MO-US019522.
PR      18-JUN-2001; 2001US-0298994P.
PR      18-JUN-2001; 2001US-0299134P.
PR      04-OCT-2001; 2001US-00972446.
PR      06-JUN-2002; 2002US-00299134.
PR      07-JUN-2002; 2002US-00298994.
XX      (CURA-) CURAGEN CORP.
XX      Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;
PI      Pena CE, Spaderna SK, Zhong M;
XX      WPI; 2003-167441/16.
XX      P-PSDB; ADJ87769.
XX      New MOLX polypeptides and polynucleotides, useful in gene therapy,
PT      particularly for treating or preventing e.g. cardiomyopathy,
PT      atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
PT      cancer, stroke or pain.
XX      Claim 8; SEQ ID NO 203; 378bp; English.
XX      The invention comprises the amino acid and coding sequences of novel G-
CC      coupled protein receptor-related (MOL) proteins. The DNA and protein
CC      sequences of the invention are useful for treating or preventing a MOL-
CC      associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
CC      associated with cell signal processing and metabolic pathway modulation,
CC      or diabetes. The DNA and protein sequences are also useful for the
CC      treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
CC      pain, hypercholesterolaemia, obesity, hyperemion, Crohn's disease,
CC      systemic lupus erythematosus, viral infections, bacterial infections,
CC      parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
CC      syndrome, Alzheimer's disease, tuberosus sclerosis, hypercalcaemia, or
CC      cerebral palsy. The present DNA sequence encodes a MOL protein of the
CC      invention.
XX      Sequence 1343 BP, 315 A, 403 C, 273 G, 352 T, 0 U, 0 Other;
SQ
XX      Query Match 99.7%; Score 1058.8; DB 10; Length 1343;
XX      Best Local Similarity 99.8%; Pred. No. 1.2e-283;
XX      Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy      1  ATGAGACACACGACGCCACCTCGACGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db      2  ATGAGACACACGACGCCACCTCGACGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61
Qy      61  TCGGCTGCGGCTGCGGCTTCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db      62  TCGGCTGCGGCTGCGGCTTCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
Qy      121  TTACAGCAAAATATCTTGAACATGATCATCTCTCCACAGCTGCTGCTGCTGCTGCTGCTGCT 180
Db      122  TTACAGCAAAATATCTTGAACATGATCATCTCTCCACAGCTGCTGCTGCTGCTGCTGCTGCT 181
Qy      181  TCTCTCTCAACTATCTCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db      182  TCTCTCTCAACTATCTCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
Qy      241  GGTGTTGAGACTCTCTGTTGAAGATTTCATCTTGAACATGAGTGGCTGAGGCTGCTGCTGCT 300
Db      242  GGTGTTGAGACTCTCTGTTGAAGATTTCATCTTGAACATGAGTGGCTGAGGCTGCTGCTGCT 301
Qy      301  GACAGATCATGAGAGTGTGGAATTTCTCATCTCACTCAACCTCCATATGATGATTACTGTA 360
Db      302  GACAGATCATGAGAGTGTGGAATTTCTCATCTCACTCAACCTCCATATGATGATTACTGTA 361
Qy      361  CCGTTAACATTTGACAGGTATATCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db      362  CCGTTAACATTTGACAGGTATATCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
Qy      421  TACCCAGCCCGCAACCCGAAAGTCATTGTAAGTGTATACATCACTGCTTCTGACCAAGC 480

```

Db 422 TACCAAGCCGACCCGGAAAGTCAATTGTATGATTTACATCACTGGCTTCCGACAGC 481

Qy 481 ATCCCTATTAATCTGTGGCCCAATCTGGAATGAGACATACATCAAGACCTCTGTGCAT 540

Db 482 ATCCCTATTAATCTGTGGCCCAATCTGGAATGAGACATACATCAAGACCTCTGTGCAT 541

Qy 541 CAGCTCTCAATCTGATCCA CTGGTCAACGGTCACTGGTGGCCCTGTCATCTTCTTC 600

Db 542 CAGCTCTCAATCTGATCCA CTGGTCAACGGTCACTGGTGGCCCTGTCATCTTCTTC 601

Qy 601 ATCTTAACATCAATCAATTTGTATCAAGCTCAGAGAGAAGCAATTTGGTCTCCGTGC 660

Db 602 ATCTTAACATCAATCAATTTGTATCAAGCTCAGAGAGAAGCAATTTGGTCTCCGTGC 661

Qy 661 TACTCCAGGGGGAAGACCAACGGCATCTTGTACCATTAACCTCCATCTTTGCAACTT 720

Db 662 TACTCCAGGGGGAAGACCAACGGCATCTTGTACCATTAACCTCCATCTTTGCAACTT 721

Qy 721 TGGGCCCCCGCATCATCATGATCTTTACCACTTATGGGGGCCCATTCAGAAACGC 780

Db 722 TGGGCCCCCGCATCATCATGATCTTTACCACTTATGGGGGCCCATTCAGAAACGC 781

Qy 781 TGGCTGTGACATCATGTGCCAATTTGCCAATGTCAATGCAAGCCCTTCTGAACAGCAATC 840

Db 782 TGGCTGTGACATCATGTGCCAATTTGCCAATGTCTAAGCTAAGCCCTTCTGAACAGCAATC 841

Qy 841 AACTTCTTCTCTACTGCTTCAATCAGCAAGCGGTTCCGACCATGAGACGGCACGCTC 900

Db 842 AACTTCTTCTCTACTGCTTCAATCAGCAAGCGGTTCCGACCATGAGACGGCACGCTC 901

Qy 901 AAGGCTTCTTCAAGGCGCAGAGAAGAACCTGTACAGTTCTACACCAATCATTAATCTTTC 960

Db 902 AAGGCTTCTTCAAGGCGCAGAGAAGAACCTGTACAGTTCTACACCAATCATTAATCTTTC 961

Qy 961 ATTAACAAGTACGCCCTGATCTTCCGCGCAAACTCACTCAATCAAGATCTGTGTAC 1020

Db 962 ATTAACAAGTACGCCCTGATCTTCCGCGCAAACTCACTCAATCAAGATCTGTGTAC 1021

Qy 1021 CAGTATGACAAAATGAGAAAACCTATAAAGATATCCCGTGA 1062

Db 1022 CAGTATGACAAAATGAGAAAACCTATAAAGATATCCCGTGA 1063

ADIT79324	ID	ADIT79324 standard; DNA; 1343 BP.
XX	AC	ADIT79324;
XX	DT	22-APR-2004 (first entry)
XX	NOV11	coding sequence, SEQ ID 21.
XX	Cytostatic;	Immunosuppressive; Antiallergic; Antimicrobial; Vasotropic;
XX	Respiratory;	Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;
XX	Ophthalmological;	Antipsoriatic; Neuroprotective; Noctropic;
XX	Antiarteriosclerotic;	Hypotensive; Cardiant; Cerebroprotective;
XX	Gene Therapy; NOVX;	human; cancer; myelogenous leukaemia;
XX	congenital neonatal autoimmune thrombocytopenia;	immunological disorder
XX	allergy; infection; asthma;	lung disease; reproductive disorder;
XX	haemangioma; deafness;	liver cirrhosis; hepatitis C; gastric disorder;
XX	diabetic retinopathy; psoriasis;	multiple sclerosis; atherosclerosis;
XX	hyperextension stroke; heart failure;	chromosome 1; NOV11;
XX	chemokine receptor; gene;	de.
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	2..1063
FT		/tag= a
XX		/product= "NOV11"

PN WO2004009635-A2.
 XX 29-JAN-2004.
 PD
 XX
 PF 04-OCT-2001; 2001WO-US031292.
 XX
 PR 20-MAR-2001; 2001US-00813432.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Taupier RJ, Majumder K, Spaderma SK, Smithson G, Mezes PS,
 PI Vermet CAM;
 XX
 DR WPI; 2004-123380/12.
 XX
 PS F-PSDB; ADI79325.
 XX
 PS Claim 8; Page 43-44; 158pp; English.
 XX
 CC The present invention relates to novel NOVX proteins and their coding
 CC sequences (ADI79304-ADI79327). The sequences are useful for the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease associated with the protein, or for diagnosing and treating
 CC disorders associated with the NOVX protein, such as cancer, myelogenous
 CC leukaemia, congenital neonatal autoimmune thrombocytopaenia,
 CC immunological disorders, allergy and infection, asthma, lung diseases,
 CC reproductive disorders, male and female reproductive diseases,
 CC haemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,
 CC hyperbolic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,
 CC hypertension, stroke and heart failure. NOVX1 represents a new member of
 CC the chemokine receptor family and the gene is located on chromosome 1.
 XX
 SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

Query Match	Best Local Similarity	99.7%;	Score 1058.8;	DB 12;	Length 1343;
Matches 1060;	Conservative	0;	Mismatches 2;	Indels	0;
		99.8%;	Pred. No. 1.2e-283;		
QY	1	ATGAGACACACCCACGCCACCTGCTGCAGACCAACAGCTGCTGTCTTGGTGGTCCCCGGAC	60		
Db	2	ATGAGACACACCCACGCCACCTGCTGCAGACCAACAGCTGCTGTCTTGGTGGTCCCCGGAC	61		
QY	61	TGCGCTTGGGCTTTGGGTTTGGTGGCGGTGGTCTACACAGCTCTTGGCTGGTGGT	120		
Db	62	TGCGCTTGGGCTTTGGGTTTGGTGGCGGTGGTCTACACAGCTCTTGGCTGGTGGT	121		
QY	121	TTACAGAGAAATCTTGACAGTATCTCTCTCCCGCTGGTGGCAGAAAGACAGAG	180		
Db	122	TTACAGAGAAATCTTGACAGTATCTCTCTCCCGCTGGTGGCAGAAAGACAGAG	181		
QY	181	TCTCTCTACAACTATCTCTTGGCACTGCTGCTGCCACATCTTGGTGGTCTTCTTCTATA	240		
Db	182	TCTCTCTACAACTATCTCTTGGCACTGCTGCTGCCACATCTTGGTGGTCTTCTTCTATA	241		
QY	241	GTGTTTGGGACCTTCTCTTGGGAAGATTTCATCTTGAACATGCAGATGCTCAGGTCCC	300		
Db	242	GTGTTTGGGACCTTCTCTTGGGAAGATTTCATCTTGAACATGCAGATGCTCAGGTCCC	301		
QY	301	GACAAAGATCATAGAAAGCTGGAATTTCTCATTCATCCACCTCCATATGATTACTGTA	360		
Db	302	GACAAAGATCATAGAAAGCTGGAATTTCTCATTCATCCACCTCCATATGATTACTGTA	361		
QY	361	CCGTTAAACATTGACAGGATATATGCTGTGTGCCACCCGCTCAAGTACACACGAGTTCA	420		
Db	362	CCGTTAAACATTGACAGGATATATGCTGTGTGCCACCCGCTCAAGTACACACGAGTTCA	421		
QY	421	TACCCAGCCCGACCCGGAAAGTCATTGTAAAGTTTACATCACTGGCTTCTGACACG	480		
Db	422	TACCCAGCCCGACCCGGAAAGTCATTGTAAAGTTTACATCACTGGCTTCTGACACG	481		
QY	481	ATCCCTATTACTGTGGCCCAATCTGGACATGGAATGGAAGTAAATGACACCTCTGGCAT	540		
Db	482	ATCCCTATTACTGTGGCCCAATCTGGACATGGAAGTAAATGACACCTCTCTGGCAT	541		

QY 541 CAGTCTCTATCTGATCCAGTGGTTCACCGTCTACCTGGTCCGTCTCATCTTCTTC 600
DB 542 CAGTCTCTATCTGATCCAGTGGTTCACCGTCTACCTGGTCCGTCTCATCTTCTTC 601
QY 601 ATCTTGAATCATCATCTTGTGTCAAGCTCAGAGAGAAAGCAATTTTCCGTCCGTGCG 660
DB 602 ATCTTGAATCATCATCTTGTGTCAAGCTCAGAGAGAAAGCAATTTTCCGTCCGTGCG 661
QY 661 TACTCCAGGGGAGAGACCAACCGCATCTTGTTCACCATTAACCTTCATCTTGGCACAATT 720
DB 662 TACTCCAGGGGAGAGACCAACCGCATCTTGTTCACCATTAACCTTCATCTTGGCACAATT 721
QY 721 TGGGCCCCCGGATCATCATGATTTCTTTTACCACTCTATGGGGCGCCATCCGAAACCGC 780
DB 722 TGGGCCCCCGGATCATCATGATTTCTTTTACCACTCTATGGGGCGCCATCCGAAACCGC 781
QY 781 TGGCTGTGATCATCATGATTCGACATTTGCCAATGCTAGCCCTTCTGAACAGAGCATC 840
DB 782 TGGCTGTGATCATCATGATTCGACATTTGCCAATGCTAGCCCTTCTGAACAGAGCATC 841
QY 841 AACTTCTCTCTACTGCTTCATCAGCAAGCGGTTCCGACATGGCAAGCGCCACGCTC 900
DB 842 AACTTCTCTCTACTGCTTCATCAGCAAGCGGTTCCGACATGGCAAGCGCCACGCTC 901
QY 901 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACGTTCTACACATCATTAATCTTTC 960
DB 902 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACGTTCTACACATCATTAATCTTTC 961
QY 961 ATTAAGATAGACCCCTGATCTCGCGGCAAACTCACACTGATCAAGATGCTGTGTAC 1020
DB 962 ATTAAGATAGACCCCTGATCTCGCGGCAAACTCACACTGATCAAGATGCTGTGTAC 1021
QY 1021 CAGTATGACAAAAATGAAAAAAGTAAAGTATCCCGGTGA 1062
DB 1022 CAGTATGACAAAAATGAAAAAAGTAAAGTATCCCGGTGA 1063
RESULT 11
AD056003
ID AD056003 standard; cDNA; 1343 BP.
XX
AC AD056003;
XX
DT 15-JUL-2004 (first entry)
XX
DE DNA encoding human NOV11.
XX
KM human; Gene; 88; cancer; obesity; diabetic nephropathy;
XX acute pancreatitis; stroke; multiple sclerosis.
XX
OS Homo sapiens.
XX
OS
XX
PN US2004058862-A1.
XX
PD 25-MAR-2004.
XX
PF 18-SEP-2002; 2002US-00246583.
XX
PR 18-SEP-2002; 2002US-00246583.
XX
PA (MAUI/) MAJIMDER K.
XX
PI MeJunder K;
XX
DR WPI; 2004-268835/25.
XX
DR P-PSDB; AD056004.
XX
PT Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic
XX nephropathy, acute pancreatitis, strokes and multiple sclerosis.
XX
PS Disclosure; Page 28; 87pp; English.
XX
CC The invention relates to novel isolated NOVX nucleic acids and encoded

CC polypeptides. The nucleic acids, polypeptides and antibodies raised
CC against the polypeptides are useful for preventing or treating diseases
CC associated with aberrant NOVX expression or activity e.g., cancer,
CC obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple
CC sclerosis. The present sequence represents a NOVX nucleic acid of the
CC invention.
XX
SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
Query Match 99.7%; Score 1058.8; DB 12; Length 1343;
Best Local Similarity 99.8%; Pred. No. 1.2e-283;
Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGACACACGCGCCCACTCCGAGCAACAGTGGTCTTGTGTGTCCTCCGCGC 60
DB 2 ATGAGACACACGCGCCCACTCCGAGCAACAGTGGTCTTGTGTGTCCTCCGCGC 61
QY 61 TGGGCTGGGCTTGGGTTTGGTGGCCGGTCTCTACTACAGGCTCTTGGTGGCTGGT 120
DB 62 TGGGCTGGGCTTGGGTTTGGTGGCCGGTCTCTACTACAGGCTCTTGGTGGCTGGT 121
QY 121 TTACGACCAATATCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAG 180
DB 122 TTACGACCAATATCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAG 181
QY 181 TCTCTCTACATATCTTGGGCACTGGTGGTGGCAAGTCTGGTCTCTTTTTCAT 240
DB 182 TCTCTCTACATATCTTGGGCACTGGTGGTGGCAAGTCTGGTCTCTTTTTCAT 241
QY 241 GTGTTTGGGCTTGGGTTTGGGAAATTCATCTTGAAGTGGTGGTGGTGGTGGTGGT 300
DB 242 GTGTTTGGGCTTGGGTTTGGGAAATTCATCTTGAAGTGGTGGTGGTGGTGGTGGT 301
QY 301 GACAGATCATAGAGTGTGGAATTTCTCATCTCAGACCTCCATATGATGATTA 360
DB 302 GACAGATCATAGAGTGTGGAATTTCTCATCTCAGACCTCCATATGATGATTA 361
QY 361 CGGTTAACATTTGACAGATATATCGGTCTGCGCACCGCTCAAGTACACAGGCTCT 420
DB 362 CGGTTAACATTTGACAGATATATCGGTCTGCGCACCGCTCAAGTACACAGGCTCT 421
QY 421 TACCGACCGGACCGGAAAGTCAATTTGAAGTTTATCATCAGCTGTCTGACAGC 480
DB 422 TACCGACCGGACCGGAAAGTCAATTTGAAGTTTATCATCAGCTGTCTGACAGC 481
QY 481 ATCCCTATTTACTGTGGCCCAACATCTGAGTGAAGACTATACATCAGACTGTGAT 540
DB 482 ATCCCTATTTACTGTGGCCCAACATCTGAGTGAAGACTATACATCAGACTGTGAT 541
QY 541 CAGTCTCTATCTGATCCAGTGGTTCACCGTCTACCTGGTCCGTCTCATCTTCTTC 600
DB 542 CAGTCTCTATCTGATCCAGTGGTTCACCGTCTACCTGGTCCGTCTCATCTTCTTC 601
QY 601 ATCTTGAATCATCATCTTGTGTCAAGCTCAGAGAGAAAGCAATTTTCCGTCCGTGCG 660
DB 602 ATCTTGAATCATCATCTTGTGTCAAGCTCAGAGAGAAAGCAATTTTCCGTCCGTGCG 661
QY 661 TACTCCAGGGGAGAGACCAACCGCATCTTGTTCACCATTAACCTTCATCTTGGCACAATT 720
DB 662 TACTCCAGGGGAGAGACCAACCGCATCTTGTTCACCATTAACCTTCATCTTGGCACAATT 721
QY 721 TGGGCCCCCGGATCATCATGATTTCTTTTACCACTCTATGGGGCGCCATCCGAAACCGC 780
DB 722 TGGGCCCCCGGATCATCATGATTTCTTTTACCACTCTATGGGGCGCCATCCGAAACCGC 781
QY 781 TGGCTGTGATCATCATGATTCGACATTTGCCAATGCTAGCCCTTCTGAACAGAGCATC 840
DB 782 TGGCTGTGATCATCATGATTCGACATTTGCCAATGCTAGCCCTTCTGAACAGAGCATC 841
QY 841 AACTTCTCTCTACTGCTTCATCAGCAAGCGGTTCCGACATGGCAAGCGCCACGCTC 900
DB 842 AACTTCTCTCTACTGCTTCATCAGCAAGCGGTTCCGACATGGCAAGCGCCACGCTC 901

QY 661 TACTCCACGGGGAAGACCAACCGCATCTTGTGACCACTTACCTTCATCTTTGGCACACTT 720
 DB 1113 TACTCCACGGGGAAGACCAACCGCATCTTGTGACCACTTACCTTCATCTTTGGCACACTT 1172
 QY 721 TGGGCCCCCGCATCATCATGATTTCTTTTACCACTTATGGGCGCCCATCCAGAACCGG 780
 DB 1173 TGGGCCCCCGCATCATCATGATTTCTTTTACCACTTATGGGCGCCCATCCAGAACCGG 1232
 QY 781 TGGCTGTGTGACATCATGTCCTGACATTTGCCAATGCTAGCCCTTCTGAAACAGAGCATC 840
 DB 1233 TGGCTGTGTGACATCATGTCCTGACATTTGCCAATGCTAGCCCTTCTGAAACAGAGCATC 1292
 QY 841 AACTTCTTCTCTCATGCTTCTCATCAGCAACCGGTTCCGACCAATGGAGCCGACAGCTC 900
 DB 1293 AACTTCTTCTCTCATGCTTCTCATCAGCAACCGGTTCCGACCAATGGAGCCGACAGCTC 1352
 QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATTAATCTTTTC 960
 DB 1353 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATTAATCTTTTC 1412
 QY 961 ATTAACAAGTAGCCCTTGATCTTGGCCGCAAACTCAACATGATGCTGTGTAC 1020
 DB 1413 ATTAACAAGTAGCCCTTGATCTTGGCCGCAAACTCAACATGATGCTGTGTAC 1472
 QY 1021 CAGTATGACAAATAATGAAAAAAGTATTAAGTATCCCGCTGA 1062
 DB 1473 CAGTATGACAAATAATGAAAAAAGTATTAAGTATCCCGCTGA 1514

RESULT 13

ABN84269
 ID ABN84269 Standard, cDNA, 1059 BP.

XX AC ABN84269,
 XX DT 23-SEP-2002 (first entry)
 XX DE Human chemokine-like receptor cDNA.
 XX KW Chemokine-like receptor; G-protein coupled receptor; receptor; human;
 KW HIV infection; cardiovascular disease; asthma;
 KW chronic obstructive pulmonary disease; cardiac; antiasthmatic;
 KW vasotropic; hypotensive; antiserhythmic; thrombolytic; anticoagulant;
 KW antiinflammatory; antiallergic; immunomodulator; gene therapy; gene; ss.
 XX OS Homo sapiens.
 XX FH Key
 FT CDS Location/Qualifiers
 FT 1..1059
 FT /tag= 8
 FT /part=1
 FT /product= "Chemokine-like receptor"
 FT /note= "the CDS does not include a stop codon"
 XX PN MO200248358-A2.
 XX PD 20-JUN-2002.
 XX PF 12-DEC-2001, 2001MO-EP014571.
 XX PR 14-DEC-2000, 2000US-0255150P.
 PR 02-APR-2001, 2001US-0280110P.
 PR 21-JUN-2001, 2001US-0299474P.
 XX PA (FARB) BAYER AG.
 XX PI Smolyer A, Zhu Z, Encinas J, Watanabe S, Okigami H;
 DR WPI, 2002-547858/58.
 DR P-PSDB, ABB79517.
 XX PT New isolated polynucleotide encoding a chemokine-like receptor

PT polypeptide for treating e.g. asthma, myocardial infarction, human
 PT immunodeficiency virus infection, or chronic obstructive pulmonary
 PT disease.

PS Claim 1; Fig 1, 114pp; English.

XX The present sequence is that of cDNA encoding a novel human chemokine-
 CC like receptor (see ANN79517) of 353 amino acids and having 7 putative
 CC transmembrane domains, consistent with the structure of a G-protein
 CC coupled receptor. Its closest human homologue is C-C chemokine receptor
 CC 3. The novel receptor is expressed at low levels in most tissues. It is
 CC expressed at a high level in phytohemagglutinin-stimulated CD8+ cells,
 CC but in none of the other immune cells tested. It may act as a receptor of
 CC chemottractant molecules on activated lymphocytes and be involved in
 CC cell trafficking and homing to sites of infection, inflammation or tissue
 CC injury. Regulation of activity of the novel receptor can therefore be
 CC used to treat cardiovascular, immunological and inflammatory diseases,
 CC including asthma and chronic obstructive pulmonary disease (COPD). The
 CC receptor may also be a target for viruses that reside in the nervous
 CC system. Regulating the binding of ligands, e.g. chemottractant molecules
 CC or virus particles, to the receptor can therefore be used to modulate the
 CC immune response to inhibit viral infections, including HIV infection. A
 CC claimed method of reducing activity of the receptor involves contacting a
 CC cell with a reagent (preferably an antibody, antisense oligonucleotide or
 CC ribozyme) to a product (preferably RNA or a polypeptide) encoded by a
 CC polynucleotide encoding the human chemokine-like receptor in vivo or in
 CC vitro. A claimed method of treating a chemokine-like receptor dysfunction
 CC related disease selected from HIV infection, a cardiovascular disorder,
 CC asthma or COPD uses a reagent that modulates a function of the receptor
 CC

SQ Sequence 1059 BP; 240 A; 342 C; 206 G; 271 T; 0 U; 0 Other;

Query Match 99.6%; Score 1057.4; DB 6; Length 1059;

Best Local Similarity 99.9%; Pred. No. 2,6e-283;

Matches 1058; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACACGACGACCCACCTCGCAGCCAAACAGCTGCTGTGTGTCCTCCCGG 60
 DB 1 ATGAGACACACGACGACCCACCTCGCAGCCAAACAGCTGCTGTGTGTCCTCCCGG 60
 QY 61 TGGGCTGTGGGCTTGGGTTTGGTGGCCGTGCTACTACAGCTCTTGTGCTCGGT 120
 DB 61 TGGGCTGTGGGCTTGGGTTTGGTGGCCGTGCTACTACAGCTCTTGTGCTCGGT 120
 QY 121 TTAACGACCAATATCTTGACAGTATCTCTCCAGCTGTGGCAAGACAGAG 180
 DB 121 TTAACGACCAATATCTTGACAGTATCTCTCCAGCTGTGGCAAGACAGAG 180
 QY 181 TCTCTTCAACATATCTTGGACATCGCTGCGGCGGACATCTGGTCTCTTTTCATA 240
 DB 181 TCTCTTCAACATATCTTGGACATCGCTGCGGCGGACATCTGGTCTCTTTTCATA 240
 QY 241 GTGTTTGTGACTTCTGTTTGGAAATTTTCATCTTGAAATGCTGAGTCCCC 300
 DB 241 GTGTTTGTGACTTCTGTTTGGAAATTTTCATCTTGAAATGCTGAGTCCCC 300
 QY 301 GACAGATCATGAAGAGTGTGGAATTTCTATCATTCACACCTCCATATGATTTAGTA 360
 DB 301 GACAGATCATGAAGAGTGTGGAATTTCTATCATTCACACCTCCATATGATTTAGTA 360
 QY 361 CGGTTAACAATTGACAGATATGCTGTGTCGACCCGCTCAAGTAGACACAGGCTCA 420
 DB 361 CGGTTAACAATTGACAGATATGCTGTGTCGACCCGCTCAAGTAGACACAGGCTCA 420
 QY 421 TACCCAGCCGACCCGGAAGTCAATTTAGTGTTTTACATCACTGCTTCTGACAC 480
 DB 421 TACCCAGCCGACCCGGAAGTCAATTTAGTGTTTTACATCACTGCTTCTGACAC 480
 QY 481 ATCCCTTATTTACTGTGTGGCCCAACATCTGAGATGAAGATTAATGACACCTCTGTGAT 540
 DB 481 ATCCCTTATTTACTGTGTGGCCCAACATCTGAGATGAAGATTAATGACACCTCTGTGAT 540
 QY 541 CAGGCTCATGTGATCCATGCTTACACCGTCTACCTGTGTCCTGCTCCATCTTCTTC 600

Db 499 CCGTTAACCATGACAGTATATCGCTGTCTGCCACCCCGCTCAAGTACACACAGGCTCTCA 558
 QY 421 TACCCAGCCCGCACCAGGAAAGTCAATTTAGTGTTTTACATCACTGCTTCTGACAGC 480
 Db 559 TACCCAGCCCGCACCAGGAAAGTCAATTTAGTGTTTTACATCACTGCTTCTGACAGC 618
 QY 481 ATCCCTATTACGTTGGGCGCCACATCTGGACGAGACTACATCAAGCACTCTGTGGCAT 540
 Db 619 ATCCCTATTACGTTGGGCGCCACATCTGGACGAGACTACATCAAGCACTCTGTGGCAT 678
 QY 541 CAGCTCTCATCTGATTCACACTGCTTCAACGCTTACCTGGTGGCTGCTCATCTTCTTC 600
 Db 679 CAGCTCTCATCTGATTCACACTGCTTCAACGCTTACAGAGTGGCTGCTCATCTTCTTC 738
 QY 601 ATTGTAACCTCAATCAATTTGTGTACAAAGCTCAGAGAGAGCAATTTTGTCTGCTGGC 660
 Db 739 ATCTTAACCTCAATCAATTTGTGTACAAAGCTCAGAGAGAGCAATTTTGTCTGCTGGC 798
 QY 661 TACTCAGGGGGAGAGACACCGCATCTTGTGACCAATTCACCTCCCATCTTTGGCAGACT 720
 Db 799 TACTCAGGGGGAGAGACACCGCATCTTGTGACCAATTCACCTCCCATCTTTGGCAGACT 858
 QY 721 TGGGCCCCCGCATCATCATGATTTCTTACACACTCTATGSGGCGCCATCCAGAACCGC 780
 Db 859 TGGGCCCCCGCATCATCATGATTTCTTACACACTCTATGSGGCGCCATCCAGAACCGC 918
 QY 781 TGGCTGTGCAATCATGTTCGACATTTGCCAATGCTGACCTTCTGAAACAGCCCATC 840
 Db 919 TGGCTGTGCAATCATGTTCGACATTTGCCAATGCTGACCTTCTGAAACAGCCCATC 978
 QY 841 AACTTTCTCTACAGCTTCAACAGAGCGGTTCCGACATGAGCGGCGCCAGCGCCAGCTC 900
 Db 979 AACTTTCTCTACAGCTTCAACAGAGCGGTTCCGACATGAGCGGCGCCAGCGCCAGCTC 1038
 QY 901 AAGGCTTTCTTCAAGTCCAGAGAACCTGTACAGTTCTACACCAATCAATCAATTTTCC 960
 Db 1039 AAGGCTTTCTTCAAGTCCAGAGAACCTGTACAGTTCTACACCAATCAATCAATTTTCC 1098
 QY 961 ATTAACAAGTAGCCCTGATCTTCCGCGCAACTCACACTGATCAAGATGCTGTGTAC 1020
 Db 1099 ATTAACAAGTAGCCCTGATCTTCCGCGCAACTCACACTGATCAAGATGCTGTGTAC 1158
 QY 1021 CAGTAGCAAAATGGAATAAAGTATTAAGTATCCCGCGA 1062
 Db 1159 CAGTAGCAAAATGGAATAAAGTATTAAGTATCCCGCGA 1200

RESULT 15
 ACA60889 standard; cDNA; 1526 BP.
 ID ACA60889, standard; cDNA; 1526 BP.
 AC ACA60889,
 XX
 DT 08-JUL-2003 (first entry)
 XX
 DE Human cDNA encoding 65507 protein.
 XX
 KW Human; ss; gene; cancer; aberrant cellular proliferation;
 KW differentiation; immune disorders; heart disorder; brain disorder;
 KW cardiovascular disorder; endothelial cell disorder; pain disorder;
 KW haemopoietic disorder; blood vessel disorder; metabolic disorder;
 KW liver disorder; platelet disorder.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 139..1200
 FT /tag= a
 FT /product= "Protein 65507"
 FT /note= "This CDS is specifically claimed in claim 1"
 XX
 XX US2003009017-A1.

PD 09-JAN-2003.
 XX
 XX 08-NOV-2001; 2001US-00012140.
 PF
 PR 08-NOV-2000; 2000US-0246768P.
 PR 08-NOV-2000; 2000US-0246772P.
 PR 15-NOV-2000; 2000US-0249185P.
 XX
 PA (LEIB/) LEIBY K R.
 PA (KAP/) KAPPELLER-LIBERMANN R.
 PA (GLUC/) GLUCKSMANN M A.
 XX
 PI Leiby KR, Kappeller-Libermann R, Glucksmann MA;
 XX
 XX WPI; 2003-428888/40.
 DR P-PSDB; AB009571.
 PT New isolated 38650, 28472, 5495, 65507, 81588 and 14354 nucleic acid
 PT molecules, useful for diagnosing, treating cancer, pain, or immune,
 PT heart, endothelial cell, hematopoietic, blood vessel, brain, metabolic
 PT and liver disorders.
 XX
 XX Claim 2; Fig 17; 90p; English.
 PS
 CC The invention relates to an isolated 38650 (encoding adenosine
 CC deaminase), 28472 (encoding a glycoproteinase), 5495 (encoding a 7
 CC transmembrane receptor), 65507, 81588 or 14354 nucleic acid molecule (or
 CC a sequence which is at least 60% identical to the six nucleic acids or
 CC their open reading frames, fragments of at least 15 nucleotides,
 CC naturally occurring variants, or a DNA insert of the plasmid deposited in
 CC with the American Type Culture Collection as Accession No. not defined in
 CC the specification, which encodes the amino acid sequence). Also included in
 CC are a host cell containing the nucleic acids (used to produce the
 CC protein), the encoded protein, an antibody that selectively binds to
 CC the polypeptide, and identifying a compound that binds to/modulates the
 CC activity of the polypeptide. The nucleic acid molecules, polypeptides and
 CC methods are useful for diagnosing, treating cancer, aberrant cellular
 CC proliferation and/or differentiation, immune disorders, heart disorders,
 CC cardiovascular disorders including endothelial cell disorders,
 CC haemopoietic disorders, blood vessel disorder, brain disorders, pain
 CC and metabolic disorders, liver disorders and platelet disorders (many
 CC examples of these disorders are given in the specification). The present
 CC sequence is the Human cDNA 65507
 XX
 SQ Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
 XX
 Query Match 98.9%; Score 1050.8; DB 8; Length 1526;
 Best Local Similarity 99.3%; Pred. No. 2.1e-281;
 Matches 1055; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ATGAGCAGACAGCAGCGCCACCTTCGAGCAACAGCTGCTGTGTTGTCCTCCCGC 60
 Db 139 ATGAGCAGACAGCAGCGCCACCTTCGAGCAACAGCTGCTGTGTTGTCCTCCCGC 198
 QY 61 TCGGCTGGGCGCTGGGCTTTCGTCGCGGCTACTAAGCTCTTGGTGGCGCTGG 120
 Db 199 TCGGCTGGGCGCTGGGCTTTCGTCGCGGCTACTAAGCTCTTGGTGGCGCTGG 258
 QY 121 TTACGAGCAATATCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAGACAAG 180
 Db 259 TTACGAGCAATATCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAGACAAG 318
 QY 181 TCTCTCAACATATCTTGGCACTGCTGCTGCCGACATCTTGTCTCTTTTCA 240
 Db 319 TCTCTCAACATATCTTGGCACTGCTGCTGCCGACATCTTGTCTCTTTTCA 378
 QY 241 GTGTTGAGCACTCTGTTGGAAGATTCACTTGAATGAGATGCTGAGGCTCC 300
 Db 379 GTGTTGAGCACTCTGTTGGAAGATTCACTTGAATGAGATGCTGAGGCTCC 438
 QY 301 GACAAGATCATAGAGTGTGGAATTTCTATCATCAACACCTCATATGATTA 360
 Db 439 GACAAGATCATAGAGTGTGGAATTTCTATCATCAACACCTCATATGATTA 498

```
QY 361 CCGTTACCATGACAGGTATATCGCTGTGCGCAACCGCTCAAGTACCAACGGTCTCA 420
Db |||||||
QY 499 CCGTTAACCATGACAGGTATATCGCTGTGCGCAACCGCTCAAGTACCAACGGTCTCA 558
Db |||||||
QY 421 TACCGAGCCCGACCCGGAAAGTCAATTGTAAGTGTTCATCACTGCTTCTTGACGAC 480
Db |||||||
QY 559 TACCGAGCCCGACCCGGAAAGTCAATTGTAAGTGTTCATCACTGCTTCTTGACGAC 618
Db |||||||
QY 481 ATCCCTTATTAAGTGTGCGCCCAACCTGGAAGTGAAGTCAATCAGCACTGTGCAAT 540
Db |||||||
QY 619 ATCCCTTATTAAGTGTGCGCCCAACCTGGAAGTGAAGTCAATCAGCACTGTGCAAT 678
Db |||||||
QY 541 CACGTCCTCATGTGATCACTGCTTCAACGCTTACCTGAGTCCCTGCTCATTTCTTC 600
Db |||||||
QY 679 CACGTCCTCATGTGATCACTGCTTCAACGCTTACCTGAGTCCCTGCTCATTTCTTC 738
Db |||||||
QY 601 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGAGAGCAATTTTGGTCCGTGGC 660
Db |||||||
QY 739 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGAGAGCAATTTTGGTCCGTGGC 798
Db |||||||
QY 661 TACTCCACGGGGAGAGCAACCGCATCTTGTTCACATTAACCTTCATCTTGGCACACTT 720
Db |||||||
QY 799 TACTCCACGGGGAGAGCAACCGCATCTTGTTCACATTAACCTTCATCTTGGCACACTT 858
Db |||||||
QY 721 TGGGCCCCCGCATCATCATGATTTCTTACCACTCTATGGGGGCCCATCCAGAACCGC 780
Db |||||||
QY 859 TGGGCCCCCGCATCATCATGATTTCTTACCACTCTATGGGGGCCCATCCAGAACCGC 918
Db |||||||
QY 781 TGGGTGTGACATCATCATGTCGACATTCGCAACATGCTAGCCCTTGTGAACAGGCCATC 840
Db |||||||
QY 919 TGGGTGTGACATCATCATGTCGACATTCGCAACATGCTAGCCCTTGTGAACAGGCCATC 978
Db |||||||
QY 841 AACTTCTTCTTACTCTTCAATCAGCAAGCGGTTCCGACATGGGACCGCCAGCTTC 900
Db |||||||
QY 979 AACTTCTTCTTACTCTTCAATCAGCAAGCGGTTCCGACATGGGACCGCCAGCTTC 1038
Db |||||||
QY 901 AAGGCTTTCTTCAAGTGCAGAAAGCAACTGTACAGTTTACACCAATCATTAACCTTTCC 960
Db |||||||
QY 1039 AAGGCTTTCTTCAAGTGCAGAAAGCAACTGTACAGTTTACACCAATCATTAACCTTTCC 1098
Db |||||||
QY 961 ATAAACAGTAGCCCTGGAATCTGGCCGGGAAACTCACTGCATCAAGATGCTGTGTAC 1020
Db |||||||
QY 1099 ATAAACAGTAGCCCTGGAATCTGGCCGGGAAACTCACTGCATCAAGATGCTGTGTAC 1158
Db |||||||
QY 1021 CAGTATGACAAATAATGAAAAACCTATMAAAGTATCCCGTGA 1062
Db |||||||
QY 1159 CAGTATGACAAATAATGAAAAACCTATMAAAGTATCCCGTGA 1200
Db |||||||
```

Search completed: December 14, 2004, 22:43:54
Job time : 500.373 sec8

RESULT 4

US-08-671-525B-1
; Sequence 1, Application US/08671525B
; Patent No. 5703220
; GENERAL INFORMATION:
; APPLICANT: Yamada, Tadataka
; APPLICANT: Gantz, Ira
; TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: US
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/671,525B
; FILING DATE: June 27, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-000853DVB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810)641-1600
; TELEFAX: (810)641-0270
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 951 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..951
; FEATURE:
; NAME/KEY: conflict
; LOCATION: order(269..270, 488, 490..491)
; OTHER INFORMATION: /note="Differs from sequence
; OTHER INFORMATION: published by Chhajlani and Wilberg in
; OTHER INFORMATION: five nucleotide and three amino acid
; OTHER INFORMATION: positions."
; FEATURE:
; NAME/KEY: conflict
; LOCATION: order(485, 488)
; OTHER INFORMATION: /note="Differs from sequence
; OTHER INFORMATION: published by Mountjoy, Robbins,
; OTHER INFORMATION: Mortrud and Cone in 2 nucleotide and 2
; OTHER INFORMATION: amino acid positions."
US-08-671-525B-1

Query Match 4.4%; Score 46.2; DB 1; Length 951;
Best Local Similarity 51.7%; Pred. No. 0.0034;

Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 285 GATGCTCAGTGTCCCGACAGATCTAAGAGCTGGAATTCTATCCATCCACACCTC 344
DB 333 GGTGCTGCGACGAGCTGGACAGATGTCTTGAAGTACGATCAGCTGAGCTGTCAG 392
QY 345 CATATGATTAAGTACCGTTAAGCATGTGACAGATATATCGCTGTGACACCGGCTCA 404

RESULT 5

US-08-672-109B-1
; Sequence 1, Application US/08672109B
; Patent No. 5710265
; GENERAL INFORMATION:
; APPLICANT: Yamada, Tadataka
; APPLICANT: Gantz, Ira
; TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: US
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,109B
; FILING DATE: June 27, 1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-000853DVC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810)641-1600
; TELEFAX: (810)641-0270
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 951 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..951
; FEATURE:
; NAME/KEY: conflict
; LOCATION: order(269..270, 488, 490..491)
; OTHER INFORMATION: /note="Differs from sequence
; OTHER INFORMATION: published by Chhajlani and Wilberg in
; OTHER INFORMATION: five nucleotide and three amino acid
; OTHER INFORMATION: positions."
; FEATURE:
; NAME/KEY: conflict
; LOCATION: order(485, 488)
; OTHER INFORMATION: /note="Differs from sequence
; OTHER INFORMATION: published by Mountjoy, Robbins,
; OTHER INFORMATION: Mortrud and Cone in 2 nucleotide and 2
; OTHER INFORMATION: amino acid positions."
US-08-672-109B-1

DB 393 CCTCTGCTTCGTGGGGCCATGCGCTGGACCGCTACATCTCATCTTTCTACGCACTGGC 452
QY 405 GTACACACAGGTCTCTATACCCAGCCCGCAACCGGAAAGTATTGTAAGTTTACATCAC 464
DB 453 CTACACAGCATCTGTACCTCGCGGGCGGCAAGCGTTGCGGCATCTGGGTGGC 512
QY 465 CTGCTTCCTGACACAGATCCCT 487
DB 513 CAGTGTGCTTCTAGCAGCGCTCT 535

Query Match 4.4%; Score 46.2; DB 1; Length 951;
Best Local Similarity 51.7%; Pred. No. 0.0034;
Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 285 GATGCTCAGTCCCCGACAGATCATAGAGTGTGGAATTTCTCATCCATCCACACCTC 344
DB 333 GGTGCTGACAGCTGACAAATGTCATTTGACCTGATCAGCTGACCTCATGCTGTCAG 392
QY 345 CATATGATTAATCTGACCTTAACTTACCATTTGACGTAATTCCTGTCGCCACCCGCTCAA 404
DB 393 CCTGTGCTTCTGGGCGCCATGCGGTGACCGCTACATCTTCATCTTCTAGCAGCTGCG 452
QY 405 GTACACACAGGTCTCATACCCAGCCGACCGGAAAGTCATTTGATGTTTACATCAC 464
DB 453 CTACACACAGCATTCGACCTGCGCGCGCGCGGCAAGCGCTTGGCCCATCTGGGTGGC 512
QY 465 CTGCTTCTGACAGCATCCCT 487
DB 513 CAGTGTGCTTCTGACAGCTCT 535

RESULT 6

US-08-842-045-1

Sequence 1, Application US/08842045
Patent No. 5817787

GENERAL INFORMATION:

APPLICANT: Yamada, Tadataka

APPLICANT: Gantz, Ira

TITLE OF INVENTION: Genes Encoding Melanocortin Receptors

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSER: Harness, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828

CITY: Bloomfield Hills

STATE: MI

COUNTRY: US

ZIP: 48303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/842,045

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Smith, Dean F.

REGISTRATION NUMBER: 36683

REFERENCE/DOCKET NUMBER: 2115-000853DVE

TELECOMMUNICATION INFORMATION:

TELEPHONE: (810)641-1600

TELEFAX: (810)641-0270

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 951 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 1..951

NAME/KEY: conflict

LOCATION: order(269..270, 488, 490..491)

OTHER INFORMATION: /note="Differs from sequence

OTHER INFORMATION: published by Chnaji and Wilberg in

OTHER INFORMATION: five nucleotide and three amino acid
positions."
FEATURE:
NAME/KEY: conflict
LOCATION: order(485, 488)
OTHER INFORMATION: /note="Differs from sequence
published by Mountjoy, Robbins,
Mortrud and Cone in 2 nucleotide and 2
amino acid positions."

Query Match 4.4%; Score 46.2; DB 1; Length 951;
Best Local Similarity 51.7%; Pred. No. 0.0034;
Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 285 GATGCTCAGTCCCCGACAGATCATAGAGTGTGGAATTTCTCATCCATCCACACCTC 344
DB 333 GGTGCTGACAGCTGACAAATGTCATTTGACCTGATCAGCTGACCTCATGCTGTCAG 392
QY 345 CATATGATTAATCTGACCTTAACTTACCATTTGACGTAATTCCTGTCGCCACCCGCTCAA 404
DB 393 CCTGTGCTTCTGGGCGCCATGCGGTGACCGCTACATCTTCATCTTCTAGCAGCTGCG 452
QY 405 GTACACACAGGTCTCATACCCAGCCGACCGGAAAGTCATTTGATGTTTACATCAC 464
DB 453 CTACACACAGCATTCGACCTGCGCGCGCGCGGCAAGCGCTTGGCCCATCTGGGTGGC 512
QY 465 CTGCTTCTGACAGCATCCCT 487
DB 513 CAGTGTGCTTCTGACAGCTCT 535

RESULT 7

US-08-842-238-1

Sequence 1, Application US/08842238
Patent No. 5869257

GENERAL INFORMATION:

APPLICANT: Yamada, Tadataka

APPLICANT: Gantz, Ira

TITLE OF INVENTION: Genes Encoding Melanocortin Receptors

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSER: Harness, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828

CITY: Bloomfield Hills

STATE: MI

COUNTRY: US

ZIP: 48303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/842,238

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Smith, Dean F.

REGISTRATION NUMBER: 36683

REFERENCE/DOCKET NUMBER: 2115-000853DVD

TELECOMMUNICATION INFORMATION:

TELEPHONE: (810)641-1600

TELEFAX: (810)641-0270

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 951 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 1..951
FEATURE:
NAME/KEY: conflict
LOCATION: order(269..270, 488, 490..491)
OTHER INFORMATION: /note="Differs from sequence
OTHER INFORMATION: published by Chhajlani and Wilberg in
OTHER INFORMATION: five nucleotide and three amino acid
OTHER INFORMATION: positions."
FEATURE:
NAME/KEY: conflict
LOCATION: order(485, 488)
OTHER INFORMATION: /note="Differs from sequence
OTHER INFORMATION: published by Mountjoy, Robbins,
OTHER INFORMATION: Mortrud and Cone in 2 nucleotide and 2
OTHER INFORMATION: amino acid positions."
US-08-842-238-1

Query Match 4.4%; Score 46.2; DB 2; Length 951;
Best Local Similarity 51.7%; Pred. No. 0.0034;
Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 285 GATGCTCAGTCCCGACAAATCATAGAGTCTGAATTTCTCATCCACACCTC 344
DB 333 GGCTGCGACAGCTGACATGTCTGACGTGATCAGCTCCATGCTGTCAG 392
QY 345 CATATGATTAATCTGACCGTTAACAATTGACAGATATATGCTGTGCCACCCGCTCAA 404
DB 393 CCTCTGCTTCTGGGGCCATCGCCGTGACCGCTACATCTTCTTACGCACTGCG 452
QY 405 GTACCAACAGGTCTCATATCCAGCCGACCCGAAAGTATGTATGTTTACATCAC 464
DB 453 CTACCAACAGATGTGACCTGCGGGGGGCGGCAAGCGTTGGCGGCATCTGGGTGGC 512
QY 465 CTGCTTCTGACCAAGCATCCCT 487
DB 513 CAGTGTGCTTTCAGCACGCTCT 535

RESULT 8
US-08-629-335B-1
Sequence 1, Application US/08629335B
Patent No. 6117975
GENERAL INFORMATION:
APPLICANT: Yamada, Tadafuka
TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Harner, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: US
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/629,335B
FILING DATE: July 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-000853DVA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600

TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 951 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 1..951
FEATURE:
NAME/KEY: conflict
LOCATION: order(269..270, 488, 490..491)
OTHER INFORMATION: /note="Differs from sequence
OTHER INFORMATION: published by Chhajlani and Wilberg in
OTHER INFORMATION: five nucleotide and three amino acid
OTHER INFORMATION: positions."
FEATURE:
NAME/KEY: conflict
LOCATION: order(485, 488)
OTHER INFORMATION: /note="Differs from sequence
OTHER INFORMATION: published by Mountjoy, Robbins,
OTHER INFORMATION: Mortrud and Cone in 2 nucleotide and 2
OTHER INFORMATION: amino acid positions."
US-08-629-335B-1

Query Match 4.4%; Score 46.2; DB 3; Length 951;
Best Local Similarity 51.7%; Pred. No. 0.0034;
Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 285 GATGCTCAGTCCCGACAAATCATAGAGTCTGAATTTCTCATCCACACCTC 344
DB 333 GGCTGCGACAGCTGACATGTCTGACGTGATCAGCTCCATGCTGTCAG 392
QY 345 CATATGATTAATCTGACCGTTAACAATTGACAGATATATGCTGTGCCACCCGCTCAA 404
DB 393 CCTCTGCTTCTGGGGCCATCGCCGTGACCGCTACATCTTCTTACGCACTGCG 452
QY 405 GTACCAACAGGTCTCATATCCAGCCGACCCGAAAGTATGTATGTTTACATCAC 464
DB 453 CTACCAACAGATGTGACCTGCGGGGGGCGGCAAGCGTTGGCGGCATCTGGGTGGC 512
QY 465 CTGCTTCTGACCAAGCATCCCT 487
DB 513 CAGTGTGCTTTCAGCACGCTCT 535

RESULT 9
US-09-868-552-1
Sequence 1, Application US/09868552
Patent No. 6693184
GENERAL INFORMATION:
APPLICANT: Howard, Andrew D.
APPLICANT: Macneil, Douglas J.
APPLICANT: Van Der Ploeg, Leonardus H. T.
TITLE OF INVENTION: DNA MOLECULES ENCODING SPICE VARIANTS
TITLE OF INVENTION: OF THE HUMAN MELANOCORTIN 1 RECEPTOR PROTEIN
FILE REFERENCE: 20367P
CURRENT APPLICATION NUMBER: US/09/868,552
CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: PCT/US99/29963
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/113,401
PRIOR FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1149

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-868-552-1

```

Query Match	4.4%	Score 46.2	DB 4	Length 1149
Best Local Similarity	51.7%	Pred. No. 0.0038		
Matches 105, Conservative	0	Mismatches 98	Indels 0	Gaps 0

QY	285	GATGCCCTCAGGTCCCGGACAGATCATYGAAGTGTGGAAATTTCATCATCAACAACCTC	344
Db	333	GGTGCTGCAGACACTGGACAAATGTCAATTGACGTGATCACTGCACACTCATGCTGTGCAG	392
QY	345	CATATGGAATTACTGTACCCTTTAACATTCAGAGATAATCGTGTCTGCCAACCCTCAA	404
Db	393	CCCTCTCTTCCTGGGGCCATGCGCGGTGSAACCGCTACATCTTCACATCTTCAAGCACATCGC	452
QY	405	GTACCAACAGGTCTCATTAACCAAGCCCGCAACCCGGAAGATCATYTGAATGTTTACATCAC	464
Db	453	TTCACACAGAGATGTGACCCCTGCGCGGGCGCGCAACCCGTTGCGGCATCTGGGTGGC	512
QY	465	CTGCTTCTCGACCAAGCATCCCTT	487
Db	513	CAGTGTGCTTTCAGACACGCTCT	535

RESULT 10
US-09-868-552-3

```

Sequence 3, Application us/09868552
Patent No. 6693184
GENERAL INFORMATION:
APPLICANT: Howard, Andrew D.
APPLICANT: Macneil, Douglas J.
APPLICANT: Van Der Ploeg, Leonardus H. T.
TITLE OF INVENTION: DNA MOLECULES ENCODING SPLICE VARIANTS
TITLE OF INVENTION: OF THE HUMAN MELANOCORTIN 1 RECEPTOR PROTEIN
FILE REFERENCE: 20367P
CURRENT APPLICATION NUMBER: US/09/868,552
CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: PCT/US99/29963
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/113,401
PRIOR FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1149
TYPE: DNA
ORGANISM: Homo sapien
IS-09-868-552-3

```

Query Match	4.4%;	Score 46.2;	DB 4;	Length 1149;
Best Local Similarity	51.7%;	Pred. No. 0.0038;		
Matches 105;	Conservative 0;	Mismatches 98;	Indels 0;	Gaps 0

QY	285	GATGGCTCAGGTCGCCCGACAAAGATATAGAAAGTGTGSAATTCATTCATCCACACCTC	344
Db	333	GGTGTGTGACGACGCTGACCAATGTTCATGTGAAGTCACCTGACAGCTTCATGCTGTCCAG	392
QY	345	CATATGATTAACGTAAACGTTAAACATTTGACAGGATATTCGCTGTCTGCCACCCGCTCAA	404
Db	393	CCTCTGTCTTCTGTGGCGCCATCGCGGTGACCGCTGACATCTCCATCTTCAACGACACTGC	452
QY	405	GTACACACGCGTCTTCATACCCAGCCCGCACCCCGGAAAGTCAATTGTAAGTTAAATACAC	464
Db	453	TTACACACGACATCGGACCTTGCCTCGCGCGCGGCGGAAAGCCTTTGCGGCATCTGGGTGC	512
QY	465	CTGCTTCTTGACACGACATCCCTT	487
Db	513	CAGTGTGCTTTCAGACAGCTCTT	535

RESULT 11
US-09-868-552-5

```

1  Sequence 5, Application US/09868552
2  Patent No. 6693184
3  GENERAL INFORMATION:
4  APPLICANT: Howard, Andrew D.
5  APPLICANT: MacNeil, Douglas J.
6  APPLICANT: Van Der Ploeg, Leonard
7  TITLE OF INVENTION: DNA MOLECULES
8  TITLE OF INVENTION: OF THE HUMAN
9  FILE REFERENCE: 20367P
10 CURRENT APPLICATION NUMBER: US/09/
11 CURRENT FILING DATE: 2001-06-18
12 PRIOR APPLICATION NUMBER: PCT/US99
13 PRIOR FILING DATE: 1999-12-16
14 PRIOR APPLICATION NUMBER: 60/113,4
15 PRIOR FILING DATE: 1998-12-23
16 NUMBER OF SEQ ID NOS: 46
17 SOFTWARE: FASESEQ for Windows Vers
18 SEQ ID NO 5
19 LENGTH: 1149
20 TYPE: DNA
21 ORGANISM: Homo sapien
22 US-09-868-552-5

```

Query Match	4.4%;	Score 46.2;	DB 4;	Length 1149;
Best Local Similarity	51.7%;	Pred. No. 0.0038;		
Matches 105; Conservative	0;	Mismatches 98;	Indels 0;	Gaps 0;

Oy	285	GATGCCCTCAGGCTCCCGACAGATCATAGAAATGCTGAAATTCATCCATCCACACCTC	344
Db	333	GGGCTCGCAGCAGCTGAGCAATGATCATTTGAGCTGATCACTGCAGCTCCAGTCTGTCCAG	392
Oy	345	CATATGGAATTAAGTACGCTTAACGTTAATGACAGGATATTCGCTCTCTGCGAACCCGGCTCAA	404
Db	393	CCCTCTCTTCTCGGGGGCCATCGCCGTGAGACCGCTCAATCTGCATCTTCTTAACGACACTGC	452
Oy	405	GTAACACACAGTCTCATATCCAGCCCGCAGCCCGGAAAGTCATTGTAAGTGTTAATCAAC	464
Db	453	CTAACACAGCAATGTGAGACCTCGCCGGCGCGGCAAGCCGTTGCGGCACTCTGGGTGGC	512
Oy	465	CTGCTTCTGACCAAGATCCCTC	487
Db	513	CAGTGTGCTTTCAGACGCTCT	535

RESULT 12

```

US-09-868-552-11
Sequence 11. Application US/09868552
Patent No 6693184
GENERAL INFORMATION:
APPLICANT: Howard, Andrew D.
APPLICANT: MacNeil, Douglas J.
APPLICANT: Van Der Ploeg, Leonardus H. T.
TITLE OF INVENTION: DNA MOLECULES ENCODING SPLICE VARIANTS
FILE REFERENCE: 20367P
CURRENT APPLICATION NUMBER: US/09/868,552
CURRENT FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: PCT/US99/29963
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/113,401
PRIOR FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO. 11
LENGTH: 1149
TYPE: DNA
ORIGIN: Homo sapien
US-09-868-552-11

```

Query Match	4.4%;	Score 46.2;	DB 4;	Length 1149;
Best Local Similarity	51.7%;	Pred. No. 0.0038;		
Matches 105, Conservative	0;	Mismatches 98;	Indels 0;	Gaps 0;

RESULT 14
US-09-868-552-25
Sequence 25, Application US/09868552
Patent No. 6693184
GENERAL INFORMATION:
APPLICANT: Howard, Andrew D.
APPLICANT: MacNeil, Douglas J.
APPLICANT: Van Der Ploeg, Leonardus H. T.
TITLE OF INVENTION: DNA MOLECULES ENCODING SPLICE VARIANTS
OF THE HUMAN MELANOCORTIN 1 RECEPTOR PROTEIN.

Query Match	4.4%	Score 46.2	DB 4	Length 1540
Best Local Similarity	51.7%	Pred. No. 0.0045		
Matches 105	Conservative	0	Mismatches 99	Indels 0
			Gaps 0	
Qy	285	GATGCTCAGTCCCGACAGATCATGATGAGTGTCTGGAATTCATCATTCACACCTC	344	
Db	333	GATGCTGAGCAGCTGACCATGTCATTTAGAGTATCACTGCAGCTCCATGCTGTCCAG	392	

QY 345 CATATGAGTAATCTGTACCGTTAACTTTGACAGGTAATATCGTGTGCAACCCCGCTCAA 404

Db 393 CCTGTGCTTCTCTGGGCGCGCTGCACCGTGAGACCGCTCATCTCATCTTTCAAGCACTGGCG 452

QY 405 GTACCA CAGGGTCTATATCCCAAGCCGACACCGGAAAGTATTTGTAAGTGTTAATCAATCAC 464

Db 453 CTACCAACAGCATCTGTACCTTGCCTGGGAGCGCGGCAAGCGCTTGGCGGACATCTGGGTGGC 512

QY 465 TTGCTTCCGACAGCATCCCGCT 487

Db 513 CAGTGTCTCTTCAAGCAAGCTCT 535

Search completed: December 15, 2004, 02:56:17
Job time : 90.3365 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 14, 2004, 22:25:49 ; Search time 542.543 Seconds
(Without alignments)
10792.424 Million cell updates/sec

Title: US-09-995-225b-15

Perfect score: 1062
Sequence: 1 atggagcagcagcagcagccca.....ctataaagctatcccgctga 1062

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 817462

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1062	100.0	1062	9	US-09-995-225-15
2	1062	100.0	1062	10	US-09-995-225-15
3	1062	100.0	1062	16	US-10-296-294A-4
4	1062	100.0	2117	17	US-10-779-104-1
5	1060.4	99.8	1062	16	US-10-296-294A-3
6	1060.4	99.8	1202	14	US-10-094-417-1
7	1058.8	99.7	1343	10	US-09-813-432-21
8	1058.8	99.7	1343	15	US-10-174-364-21
9	1058.8	99.7	1343	16	US-10-246-583-21
10	1058.8	99.7	1343	17	US-10-689-832-21
11	1057.4	99.6	1059	16	US-10-450-590-1
12	1050.8	98.9	1062	14	US-10-012-140-12

13	1050.8	98.9	1526	14	US-10-012-140-10	Sequence 10, Appl
14	1048.4	98.7	1119	10	US-09-813-432-19	Sequence 19, Appl
15	1048.4	98.7	1119	15	US-10-174-364-19	Sequence 19, Appl
16	1048.4	98.7	1119	16	US-10-246-583-19	Sequence 19, Appl
17	1048.4	98.7	1119	17	US-10-689-832-19	Sequence 19, Appl
18	1048.4	98.7	1130	16	US-10-333-946-25	Sequence 25, Appl
19	1048.4	98.7	2198	15	US-10-314-076-1	Sequence 1, Appl
20	1039.8	97.9	1119	15	US-10-314-076-3	Sequence 3, Appl
21	1039.8	97.9	2189	14	US-10-219-834-6	Sequence 6, Appl
22	945.2	89.0	1062	16	US-10-450-590-9	Sequence 9, Appl
23	940.2	88.5	957	15	US-10-174-364-84	Sequence 84, Appl
24	940.2	88.5	957	16	US-10-246-583-84	Sequence 84, Appl
25	936.4	88.2	1002	15	US-10-079-384-25	Sequence 25, Appl
26	936.4	88.2	1158	16	US-10-343-650A-21	Sequence 21, Appl
27	936	88.1	1032	16	US-10-450-590-5	Sequence 5, Appl
28	936	88.1	1070	16	US-10-450-590-6	Sequence 4, Appl
29	936	88.1	1826	16	US-10-450-590-6	Sequence 6, Appl
30	844	79.5	1038	14	US-10-094-417-19	Sequence 19, Appl
31	842.6	79.3	864	10	US-09-791-932-6	Sequence 6, Appl
32	758	71.4	795	15	US-10-225-567A-586	Sequence 586, App
33	524	49.3	930	10	US-09-791-932-3	Sequence 3, Appl
34	268.2	25.3	1466	16	US-10-328-916-4	Sequence 4, Appl
35	268.2	25.3	1962	16	US-10-333-946-29	Sequence 29, Appl
36	266.6	25.1	1125	9	US-09-995-225-5	Sequence 5, Appl
37	266.6	25.1	1125	10	US-09-995-225-5	Sequence 10, Appl
38	266.6	25.1	1125	14	US-10-012-140-15	Sequence 15, Appl
39	266.6	25.1	1427	13	US-10-094-417-13	Sequence 13, Appl
40	266.6	25.1	1530	14	US-10-011-147-1	Sequence 1, Appl
41	266.6	25.1	1719	14	US-10-012-140-13	Sequence 13, Appl
42	266.6	25.1	1793	16	US-10-328-916-3	Sequence 3, Appl
43	243.8	23.0	34118	15	US-10-017-161-1071	Sequence 1071, Ap
44	243.8	23.0	34118	15	US-10-292-788-999	Sequence 909, App
45	80	7.5	80	14	US-10-219-834-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1
US-09-995-225-15
; Sequence 15, Application US/09995225
; Publication No. US20020193584A1
GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pride, Cameron
; TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 1062
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: No. US20020193584A1 Sequence
US-09-995-225-15

Query Match 100.0%; Score 1062; DB 9; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2e-310;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGACACGACGCGCCACCTGCGACGCCAAGCTGCTGTGTGGTCCCGGCGC 60
DB 1 ATGAGACACGACGCGCCACCTGCGACGCCAAGCTGCTGTGTGGTCCCGGCGC 60
QY 61 TCGGCGCTGGCGCTGGCTTTGCGTCCGCTGCTTACTACAGCCTCTTGTGCTGCTGCT 120
DB 61 TCGGCGCTGGCGCTGGCTTTGCGTCCGCTGCTTACTACAGCCTCTTGTGCTGCTGCTG 120
QY 121 TTACACGCAATAATCTTGACAGTGCATCTCTCCAGCTGGTGGCAAGAACAGAAC 180
DB 121 TTACACGCAATAATCTTGACAGTGCATCTCTCCAGCTGGTGGCAAGAACAGAAC 180
QY 181 TCCCTCCACAACTATCTCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 TCCCTCCACAACTATCTCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 GTGTTGTGGAATCTCTGTTGGAAGATTGATCTTGAAGATGAGATGCTGCTGCTGCT 300
DB 241 GTGTTGTGGAATCTCTGTTGGAAGATTGATCTTGAAGATGAGATGCTGCTGCTGCT 300
QY 301 GACAAAGTCAATAGAGTGTGGAATCTCATTCATCCACAGCTCCATGATGATTAAGT 360
DB 301 GACAAAGTCAATAGAGTGTGGAATCTCATTCATCCACAGCTCCATGATGATTAAGT 360
QY 361 CCGTTAACATTAAGAGTATATCGGTCTGCGACCCGCTGATGATGATTAAGT 420
DB 361 CCGTTAACATTAAGAGTATATCGGTCTGCGACCCGCTGATGATGATTAAGT 420
QY 421 TACCCAGCCCGCACCCGAAAGTCAATGATGATTAAGTATGATGATGATGATGAT 480
DB 421 TACCCAGCCCGCACCCGAAAGTCAATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 ATCCCTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 ATCCCTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 CAGCTCTCATCTGATCACTGCTTACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 CAGCTCTCATCTGATCACTGCTTACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 ATCTTGAATCAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 ATCTTGAATCAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TACTCAGGAGGAGACACGCGCATCTGTTGATCACTTACCTGATCTTTGGACACTT 720
DB 661 TACTCAGGAGGAGACACGCGCATCTGTTGATCACTTACCTGATCTTTGGACACTT 720
QY 721 TGGGCGCCCGCGATCATGATCTTTACCACTCTATGAGGCGCCCATCAAGACCGC 780
DB 721 TGGGCGCCCGCGATCATGATCTTTACCACTCTATGAGGCGCCCATCAAGACCGC 780
QY 781 TGGGCGGAGCAGTATGCTGCACTTGGCAAGTCTGAGCCCTTGTGAACAGCCATC 840
DB 781 TGGGCGGAGCAGTATGCTGCACTTGGCAAGTCTGAGCCCTTGTGAACAGCCATC 840
```

```
QY 841 AACTCTCTCTACTGCTTCATCAGCAGCGGTTCCGACCATGCGACCGCCACGCTC 900
DB 841 AACTCTCTCTCTACTGCTTCATCAGCAGCGGTTCCGACCATGCGACCGCCACGCTC 900
QY 901 AAGGCTTTCTTCAAGTCCAGAGCAAGCACTGATGATGATGATGATGATGATGATGAT 960
DB 901 AAGGCTTTCTTCAAGTCCAGAGCAAGCACTGATGATGATGATGATGATGATGATGAT 960
QY 961 ATACAGTAAAGCCCTCGATCTGCGCGGAACCTCACTGATCAAGATGCTGTGATC 1020
DB 961 ATACAGTAAAGCCCTCGATCTGCGCGGAACCTCACTGATCAAGATGCTGTGATC 1020
QY 1021 CAGTATGACAAATAATGAAAAACCTATTAAGATCCCGTGA 1062
DB 1021 CAGTATGACAAATAATGAAAAACCTATTAAGATCCCGTGA 1062
```

RESULT 2
US-09-995-225-15
Sequence 15, Application US/09995225
Publication No. US20030139588A9
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Chu, Zhi Liang
APPLICANT: Dang, Huang T.
APPLICANT: Lowitz, Kevin P.
APPLICANT: Pride, Cameron
TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
FILE REFERENCE: AREN-0308
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 1062
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: No. US20030139588A9e1 Sequence
US-09-995-225-15

Query Match 100.0%; Score 1062; DB 10; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2e-310;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACACGACGCGCCACCTGCGACGCCAAGCTGCTGTGTGGTCCCGGCGC 60

Db 1 ATGAGACACGACGCGCCACCTCGGACGCCAAGCTCGCTGCTTGTGTGTCGCCGCGC 60
QY 61 TCGGCTCGGCGCTTGGGTTTGTGCGCCGTGTCTACTACAGCTCTTGTGTCCTCGGT 120
Db 61 TCGGCTCGGCGCTTGGGTTTGTGCGCCGTGTCTACTACAGCTCTTGTGTCCTCGGT 120
QY 121 TTACGACGAAATATCTTGAAGATCATCTCTCCAGCTGTGTGCGAAGACGAG 180
Db 121 TTACGACGAAATATCTTGAAGATCATCTCTCCAGCTGTGTGCGAAGACGAG 180
QY 181 TCCTCTACAACTATCTTGTGCACTCGCTGCGCAACATCTTGTGCTCTTTTCTA 240
Db 181 TCCTCTACAACTATCTTGTGCACTCGCTGCGCAACATCTTGTGCTCTTTTCTA 240
QY 241 GTGTTGTGACCTTCTGTGGAAGATTTCACTTGAACATGACAGATGCTCAGGTCCC 300
Db 241 GTGTTGTGACCTTCTGTGGAAGATTTCACTTGAACATGACAGATGCTCAGGTCCC 300
QY 301 GACAAATCATTAAGCTGGAATTCTCATTCATCCACACCTCCCAATGGAATTACTGTA 360
Db 301 GACAAATCATTAAGCTGGAATTCTCATTCATCCACACCTCCCAATGGAATTACTGTA 360
QY 361 CGGTAAACATTAAGCAAGTATATGCTGTGCGCACCGGCTCAAGTACACACGATCTCA 420
Db 361 CGGTAAACATTAAGCAAGTATATGCTGTGCGCACCGGCTCAAGTACACACGATCTCA 420
QY 421 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTAATCACTCTGCTTCTGACACG 480
Db 421 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTAATCACTCTGCTTCTGACACG 480
QY 481 ATCCCTATTAATCTGTGTGCGCCCAATCTGGAAGTGAAGTCAATCAAGCACTCTGAGAT 540
Db 481 ATCCCTATTAATCTGTGTGCGCCCAATCTGGAAGTGAAGTCAATCAAGCACTCTGAGAT 540
QY 541 CAGCTCTCATCTGATTCACCTGCTTCAACGCTCAGAGAGAGCAATTTTCCGTCCGTCGC 600
Db 541 CAGCTCTCATCTGATTCACCTGCTTCAACGCTCAGAGAGAGCAATTTTCCGTCCGTCGC 600
QY 601 ATCTGAATCATCATCTTGTGTAAGTCTCAGAGAGAGCAATTTTCCGTCCGTCGC 660
Db 601 ATCTGAATCATCATCTTGTGTAAGTCTCAGAGAGAGCAATTTTCCGTCCGTCGC 660
QY 661 TACTCAACGCGGGAAGCAACCGGCACTTGTTCACCACTTCACTCTTGTGCAACACT 720
Db 661 TACTCAACGCGGGAAGCAACCGGCACTTGTTCACCACTTCACTCTTGTGCAACACT 720
QY 721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGAGGCGCCATCAGAAACGC 780
Db 721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGAGGCGCCATCAGAAACGC 780
QY 781 TGGCTGTGTCATCATCTGTCGACATTTGCCAATGCTAGCCCTTGTGAACAGACCATC 840
Db 781 TGGCTGTGTCATCATCTGTCGACATTTGCCAATGCTAGCCCTTGTGAACAGACCATC 840
QY 841 AACTTCTCTCTCATCTGTCATCAGCAAGCGGTTCCGCACTATGGAAGCGCCACAGCTC 900
Db 841 AACTTCTCTCTCATCTGTCATCAGCAAGCGGTTCCGCACTATGGAAGCGCCACAGCTC 900
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTCTTACACCAATCTAATCTTTCC 960
Db 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTCTTACACCAATCTAATCTTTCC 960
QY 961 ATTAACAAGTAGCCCTGTGATCTGCGCGGCAAACTCACTGCAATCAAGATGCTGTATC 1020
Db 961 ATTAACAAGTAGCCCTGTGATCTGCGCGGCAAACTCACTGCAATCAAGATGCTGTATC 1020
QY 1021 CAGTATGACAAATAATGAAAACTATTAAGTATCCCGTGA 1062
Db 1021 CAGTATGACAAATAATGAAAACTATTAAGTATCCCGTGA 1062

RESULT 3
US-10-296-294A-4

; Sequence 4, Application US/10296294A
; Publication No. US20040029224A1
; GENERAL INFORMATION:
; APPLICANT: TERAO, Yasuko
; APPLICANT: MATSUI, Hideki
; APPLICANT: SHINTANI, Yasushi
; TITLE OF INVENTION: No. US20040029224A1e1 G Protein-Coupled Receptor and its DNA
; FILE REFERENCE: 2734 USJP
; CURRENT APPLICATION NUMBER: US/10/296, 294A
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: PCT/JP01/04643
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: JP 2000-170446
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: JP 2000-194926
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Human
US-10-296-294A-4
Query Match 100.0%; Score 1062; DB 16; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2e-310;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACACGACGCGCCACCTCGGACGCCAAGCTCGCTGCTTGTGTGTCGCCGCGC 60
Db 1 ATGAGACACGACGCGCCACCTCGGACGCCAAGCTCGCTGCTTGTGTGTCGCCGCGC 60
QY 61 TCGGCTCGGCGCTTGGGTTTGTGCGCCGTGTCTACTACAGCTCTTGTGTCCTCGGT 120
Db 61 TCGGCTCGGCGCTTGGGTTTGTGCGCCGTGTCTACTACAGCTCTTGTGTCCTCGGT 120
QY 121 TTACGACGAAATATCTTGAAGATCATCTCTCCAGCTGTGTGCGAAGACGAG 180
Db 121 TTACGACGAAATATCTTGAAGATCATCTCTCCAGCTGTGTGCGAAGACGAG 180
QY 181 TCCTCTACAACTATCTTGTGCACTCGCTGCGCAACATCTTGTGCTCTTTTCTA 240
Db 181 TCCTCTACAACTATCTTGTGCACTCGCTGCGCAACATCTTGTGCTCTTTTCTA 240
QY 241 GTGTTGTGACCTTCTGTGGAAGATTTCACTTGAACATGACAGATGCTCAGGTCCC 300
Db 241 GTGTTGTGACCTTCTGTGGAAGATTTCACTTGAACATGACAGATGCTCAGGTCCC 300
QY 301 GACAAATCATTAAGCTGGAATTCTCATTCATCCACACCTCCCAATGGAATTACTGTA 360
Db 301 GACAAATCATTAAGCTGGAATTCTCATTCATCCACACCTCCCAATGGAATTACTGTA 360
QY 361 CGGTAAACATTAAGCAAGTATATGCTGTGCGCACCGGCTCAAGTACACACGATCTCA 420
Db 361 CGGTAAACATTAAGCAAGTATATGCTGTGCGCACCGGCTCAAGTACACACGATCTCA 420
QY 421 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTAATCACTCTGCTTCTGACACG 480
Db 421 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTAATCACTCTGCTTCTGACACG 480
QY 481 ATCCCTATTAATCTGTGTGCGCCCAATCTGGAAGTGAAGTCAATCAAGCACTCTGAGAT 540
Db 481 ATCCCTATTAATCTGTGTGCGCCCAATCTGGAAGTGAAGTCAATCAAGCACTCTGAGAT 540
QY 541 CAGCTCTCATCTGATTCACCTGCTTCAACGCTCAGAGAGAGCAATTTTCCGTCCGTCGC 600
Db 541 CAGCTCTCATCTGATTCACCTGCTTCAACGCTCAGAGAGAGCAATTTTCCGTCCGTCGC 600
QY 601 ATCTGAATCATCATCTTGTGTAAGTCTCAGAGAGAGCAATTTTCCGTCCGTCGC 660
Db 601 ATCTGAATCATCATCTTGTGTAAGTCTCAGAGAGAGCAATTTTCCGTCCGTCGC 660
QY 661 TACTCAACGCGGGAAGCAACCGGCACTTGTTCACCACTTCACTCTTGTGCAACACT 720
Db 661 TACTCAACGCGGGAAGCAACCGGCACTTGTTCACCACTTCACTCTTGTGCAACACT 720

Db 661 TACTCAGCGGGAGAACACCGCCATCTTTTCACATTAACCTCCATCTTTGCCACACTT 720
QY 721 TGGGCCCCCGGCGATCATCATGATTTCTTTACCACTTATGAGGCGCCCATCCAGAACGCG 780
Db 721 TGGGCCCCCGGCGATCATCATGATTTCTTTACCACTTATGAGGCGCCCATCCAGAACGCG 780
QY 781 TGGGCGTGTGACATCATGATTCGCAATTTGCAACATGCTGAGCGCTTCTGAACCAAGCTTC 840
Db 781 TGGGCGTGTGACATCATGATTCGCAACATGCTGAGCGCTTCTGAACCAAGCTTC 840
QY 841 AACTTCTGCTCATCTGCTTTCATGACAAAGCGGTTCCGACCATGAGCGCGCAAGCTTC 900
Db 841 AACTTCTGCTCATCTGCTTTCATGACAAAGCGGTTCCGACCATGAGCGCGCAAGCTTC 900
QY 901 AAGGCTTTCTTCAAGTGTGACAGAACCTGTGACAGTTCACCAATCATTAATCTTTTCC 960
Db 901 AAGGCTTTCTTCAAGTGTGACAGAACCTGTGACAGTTCACCAATCATTAATCTTTTCC 960
QY 961 ATACCAAGTAGAGCCCTGATCTGCGCGCAAACTGACATGACATCAAGATGCTGTGTAC 1020
Db 961 ATACCAAGTAGAGCCCTGATCTGCGCGCAAACTGACATGACATCAAGATGCTGTGTAC 1020
QY 1021 CAGTATGACAAAATGAAAACCTATAAAGTATCCCGGTGA 1062
Db 1021 CAGTATGACAAAATGAAAACCTATAAAGTATCCCGGTGA 1062

RESULT 4

US-10-779-104-1
; Sequence 1, Application US/10779104
; Publication No. US20040161799A1
; GENERAL INFORMATION:
; APPLICANT: Susan Croll-Murphy
; APPLICANT: Andrew J. Murphy
; TITLE OF INVENTION: KOR3L-like-Proteins and Methods of Modulating KOR3L-Mediated Activi
; FILE REFERENCE: RSG 1000A
; CURRENT APPLICATION NUMBER: US/10/779,104
; PRIOR APPLICATION NUMBER: 2004-02-13
; PRIOR FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/447,447
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-779-104-1

Query Match 100.0%; Score 1062; DB 17; Length 2117;
Best Local Similarity 100.0%; Pred. No. 2,8e-310;

Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCAGACGACGCGCCACCTCGACGCAACAGCTGCTGTGTTGGTGTCCCGCGGC 60
Db 280 ATGGAGCAGACGACGCGCCACCTCGACGCAACAGCTGCTGTGTTGGTGTCCCGCGGC 339
QY 61 TCGGCTGTGCGGCTTGGGTTTCGTGCGCGGCTGCTCACTCAAGCTTGTGCTGCTGCT 120
Db 340 TCGGCTGTGCGGCTTGGGTTTCGTGCGCGGCTGCTCACTCAAGCTTGTGCTGCTGCT 399
QY 121 TTACCGCAAAATATTTTGAAGATGATCTCTTCCAGCTGTGTGCAAGAAAGACAGAG 180
Db 400 TTACCGCAAAATATTTTGAAGATGATCTCTTCCAGCTGTGTGCAAGAAAGACAGAG 459
QY 181 TCTCTCTCAACATATCTTGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 460 TCTCTCTCAACATATCTTGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
QY 241 GTGTTGTGAATCTTGTGGAAGATTTTCACTTGAACAGAGATGCTGCTGCTGCTGCT 300
Db 520 GTGTTGTGAATCTTGTGGAAGATTTTCACTTGAACAGAGATGCTGCTGCTGCTGCTGCT 579

QY 301 GACAGATCATTAAGTGTGATTTCTATTCATCAACCTTCATATGATTAATGATTA 360
Db 580 GACAGATCATTAAGTGTGATTTCTATTCATCAACCTTCATATGATTAATGATTA 639
QY 361 CCGTTAAACATTTGACAGATTAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 640 CCGTTAAACATTTGACAGATTAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 699
QY 421 TACCCAGCCCGCAACCGGAAAGTCAATTTGATGTTTATCATCACTGCTTCTGCAAGC 480
Db 700 TACCCAGCCCGCAACCGGAAAGTCAATTTGATGTTTATCATCACTGCTTCTGCAAGC 759
QY 481 ATCCCTATTTATGCTGAGCCCAACATCTGATCTGAAGCTTATCATCAAGCACTGCTGAT 540
Db 760 ATCCCTATTTATGCTGAGCCCAACATCTGATCTGAAGCTTATCATCAAGCACTGCTGAT 819
QY 541 CAGCTCTCATCTGATCACTGCTTCAACGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 820 CAGCTCTCATCTGATCACTGCTTCAACGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 879
QY 601 ATCTGAATCAATGATTTGTTGATCAAGCTCAAGAGAGAAATTTTCTGCTGCTGCTGCT 660
Db 880 ATCTGAATCAATGATTTGTTGATCAAGCTCAAGAGAGAAATTTTCTGCTGCTGCTGCT 939
QY 661 TACTCAGCGGGAGAACACCGCCATCTTGTTCACATTAACCTTCATCTTTGCCAGCTT 720
Db 940 TACTCAGCGGGAGAACACCGCCATCTTGTTCACATTAACCTTCATCTTTGCCAGCTT 999
QY 721 TGGGCCCCCGGCGATCATCATGATTTCTTTACCACTTATGAGGCGCCCATCCAGAACGCG 780
Db 1000 TGGGCCCCCGGCGATCATCATGATTTCTTTACCACTTATGAGGCGCCCATCCAGAACGCG 1059
QY 781 TGGGCGTGTGACATCATGATTCGCAATTTGCAAGATGCTGAGCGCTTGTGAACAGAGCATTC 840
Db 1060 TGGGCGTGTGACATCATGATTCGCAAGATGCTGAGCGCTTGTGAACAGAGCATTC 1119
QY 841 AACTTCTGCTCATCTGCTTTCATGACAAAGCGGTTCCGACCATGAGCGCGCAAGCTTC 900
Db 1120 AACTTCTGCTCATCTGCTTTCATGACAAAGCGGTTCCGACCATGAGCGCGCAAGCTTC 1179
QY 901 AAGGCTTTCTTCAAGTGTGACAGAACCTGTGACAGTTCACCAATCATTAATCTTTTCC 960
Db 1180 AAGGCTTTCTTCAAGTGTGACAGAACCTGTGACAGTTCACCAATCATTAATCTTTTCC 1239
QY 961 ATACCAAGTAGAGCCCTGATCTGCGCGCAAACTGACATGACATCAAGATGCTGTGTAC 1020
Db 1240 ATACCAAGTAGAGCCCTGATCTGCGCGCAAACTGACATGACATCAAGATGCTGTGTAC 1299
QY 1021 CAGTATGACAAAATGAAAACCTATAAAGTATCCCGGTGA 1062
Db 1300 CAGTATGACAAAATGAAAACCTATAAAGTATCCCGGTGA 1341

RESULT 5

US-10-296-294A-3
; Sequence 3, Application US/10296294A
; Publication No. US20040029224A1
; GENERAL INFORMATION:
; APPLICANT: TERAO, Yasuko
; APPLICANT: MATSUI, Hideki
; TITLE OF INVENTION: SHINTANI, Yasushi
; FILE REFERENCE: 2734 USOP
; CURRENT APPLICATION NUMBER: US/10/296,294A
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: PCT/JP01/04643
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: JP 2000-170446
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: JP 2000-194926
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 8

SEQ ID NO 3
LENGTH: 1062
TYPE: DNA
ORGANISM: Human
US-10-296-294A-3

Query Match 99.8%; Score 1060.4; DB 16; Length 1062;
Best Local Similarity 99.9%; Pred. No. 6e-310; Indels 0; Gaps 0;
Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGCACAGCAGCCCACTCGAGCCCAAGCTCGCTCTTGTGTGTCCTCCCGC 60
DB 1 ATGAGCACAGCAGCCCACTCGAGCCCAAGCTCGCTCTTGTGTGTCCTCCCGC 60
QY 61 TCGGCTGCGGCTTGGGTTTCTGCGCGGTGTCTACTACAGCTCTTGTGCTCGGT 120
DB 61 TCGGCTGCGGCTTGGGTTTCTGCGCGGTGTCTACTACAGCTCTTGTGCTCGGT 120
QY 121 TTACGAGCAAAATCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAG 180
DB 121 TTACGAGCAAAATCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAG 180
QY 181 TCCTCTACAACTATCTTGGCACTCGCTGTCGCGACATCTTGTCTCTTTTCATA 240
DB 181 TCCTCTACAACTATCTTGGCACTCGCTGTCGCGACATCTTGTCTCTTTTCATA 240
QY 241 GTGTTGTGAGCTTCTGTTGGAAGTTTCACTTGAACATGCAATGCTTCAGTCCC 300
DB 241 GTGTTGTGAGCTTCTGTTGGAAGTTTCACTTGAACATGCAATGCTTCAGTCCC 300
QY 301 GACAGATCATAGAGTGTGAATCTCATTCATCCACCTCCATATGATTAATCTGTA 360
DB 301 GACAGATCATAGAGTGTGAATCTCATTCATCCACCTCCATATGATTAATCTGTA 360
QY 361 CGGTTAACATTTAGAGATATATGCTGTGTCACCGCTCAAGTACCAAGCTCTCA 420
DB 361 CGGTTAACATTTAGAGATATATGCTGTGTCACCGCTCAAGTACCAAGCTCTCA 420
QY 421 TACCCAGCCCGCAGCGGAAAGTATGTAAGTGTATCACTCTGCTTCGACACAG 480
DB 421 TACCCAGCCCGCAGCGGAAAGTATGTAAGTGTATCACTCTGCTTCGACACAG 480
QY 481 ATCCCTATTAATGAGTGTGACCAATCTGGAATGAGAGATCAATGAGACCTGTGAT 540
DB 481 ATCCCTATTAATGAGTGTGACCAATCTGGAATGAGAGATCAATGAGACCTGTGAT 540
QY 541 CAGTCTCATCTGATTCACCTGCTTCACTGCTGCTGCTGCTCATCTTCTTC 600
DB 541 CAGTCTCATCTGATTCACCTGCTTCACTGCTGCTGCTGCTCATCTTCTTC 600
QY 601 ATCTTGAATCATCTTGTGTCAAGCTCAGAGAGAGAGCAATTTGTCTCGGTGC 660
DB 601 ATCTTGAATCATCTTGTGTCAAGCTCAGAGAGAGAGCAATTTGTCTCGGTGC 660
QY 661 TACTCAGGAGGAGAGCAAGCCATCTGTTTACCATTAACCTTCATCTTGTGCAACT 720
DB 661 TACTCAGGAGGAGAGCAAGCCATCTGTTTACCATTAACCTTCATCTTGTGCAACT 720
QY 721 TGGGCCCCCGCATCATGATTTCTTACCACTTATGAGGCGGCCATCCAGAACCG 780
DB 721 TGGGCCCCCGCATCATGATTTCTTACCACTTATGAGGCGGCCATCCAGAACCG 780
QY 781 TGGCTGTGCAATCATGTCGCAATTTGCAATGCTTGTGCAATGCTTGTGCAAT 840
DB 781 TGGCTGTGCAATCATGTCGCAATTTGCAATGCTTGTGCAATGCTTGTGCAAT 840
QY 841 AACTTCTCTCTCATCTGCTTCACTCAGCAAGCGGTTCGCAATGAGCGGCAAGCTC 900
DB 841 AACTTCTCTCTCATCTGCTTCACTCAGCAAGCGGTTCGCAATGAGCGGCAAGCTC 900
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTAAGTTTCAACCAATCAATCTTTCC 960
DB 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTAAGTTTCAACCAATCAATCTTTCC 960
```

```
QY 961 ATACAGATAGCCCTTGTGATCTCCCGGCAAACTGACATGATGAATGTGCTGCTAC 1020
DB 961 ATACAGATAGCCCTTGTGATCTCCCGGCAAACTGACATGATGAATGTGCTGCTAC 1020
QY 1021 CAGTATGACAAAATGAGAAACCTATTAAGATATCCCGCTGA 1062
DB 1021 CAGTATGACAAAATGAGAAACCTATTAAGATATCCCGCTGA 1062
```

RESULT 6

US-10-094-417-1
Sequence 1, Application US/10094417
Publication No. US20030045685A1
GENERAL INFORMATION:
APPLICANT: Tian, Hui
APPLICANT: Zhao, Jiaqiang
APPLICANT: Chen, Jin-Long
APPLICANT: Cutler, Gene
APPLICANT: Tularik Inc.
TITLE OF INVENTION: No. US20030045685A1e1 Receptors
FILE REFERENCE: 018781-008110US
CURRENT APPLICATION NUMBER: US/10/094,417
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/802,803
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: US 60/276,649
PRIOR FILING DATE: 2001-03-16
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1
LENGTH: 1202
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
NAME/KEY: CDS
LOCATION: (68)..(1129)
OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
US-10-094-417-1

Query Match 99.8%; Score 1060.4; DB 14; Length 1202;
Best Local Similarity 99.9%; Pred. No. 6.4e-310; Indels 0; Gaps 0;
Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGCACAGCAGCCCACTCGAGCCCAAGCTCGCTCTTGTGTGTCCTCCCGC 60
DB 68 ATGAGCACAGCAGCCCACTCGAGCCCAAGCTCGCTCTTGTGTGTCCTCCCGC 127
QY 61 TCGGCTGCGGCTTGGGTTTCTGCGCGGTGTCTACTACAGCTCTTGTGCTCGGT 120
DB 128 TCGGCTGCGGCTTGGGTTTCTGCGCGGTGTCTACTACAGCTCTTGTGCTCGGT 187
QY 121 TTACGAGCAAAATCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAGAG 180
DB 188 TTACGAGCAAAATCTTGAAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAGAG 247
QY 181 TCCTCTACAACTATCTTGGCACTCGCTGTCGCGACATCTTGTGCTCTTTTCATA 240
DB 248 TCCTCTACAACTATCTTGGCACTCGCTGTCGCGACATCTTGTGCTCTTTTCATA 307
QY 241 GTGTTGTGAGCTTCTGTTGGAAGTTTCACTTGAACATGCAATGCTTGTGCTCGGT 300
DB 308 GTGTTGTGAGCTTCTGTTGGAAGTTTCACTTGAACATGCAATGCTTGTGCTCGGT 367
QY 301 GACAGATCATAGAGTGTGAATTTCTCATCTCACTCACTCCATATGATTAATCTGTA 360
DB 368 GACAGATCATAGAGTGTGAATTTCTCATCTCACTCACTCCATATGATTAATCTGTA 427
QY 361 CGGTTAACATTTAGAGATATATGCTGTGTCACCGCTCAAGTATCACACGCTTCA 420
DB 428 CGGTTAACATTTAGAGATATATGCTGTGTCACCGCTCAAGTATCACACGCTTCA 487
```


QY 781 TGGCTGTGACATCATGTGTCACATTGGCCAAATGCTAGCCCTTCTGAAACAGAGCATC 840
DB 782 TGGCTGTGACATCATGTGTCACATTGGCCAAATGCTAGCCCTTCTGAAACAGAGCATC 841
QY 841 AACTCTTCTCTACTGCTTCTCATCAGCAAGCGGTTCGCGACATGAGCGCGCAGCTC 900
DB 842 AACTCTTCTCTACTGCTTCTCATCAGCAAGCGGTTCGCGACATGAGCGCGCAGCTC 901
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACCAATCTTAACCTTTTTC 960
DB 902 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACCAATCTTAACCTTTTTC 961
QY 961 ATACAGAGTACCCCTGATCTGCGCGGCAACTCATCTGATCTCAAGATGCTGTATC 1020
DB 962 ATACAGAGTACCCCTGATCTGCGCGGCAACTCATCTGATCTCAAGATGCTGTATC 1021
QY 1021 CAGTATGACAAATAAGAAAACTATATAAGTATCCCGGTGA 1062
DB 1022 CAGTATGACAAATAAGAAAACTATATAAGTATCCCGGTGA 1063

RESULT 8
US-10-174-364-21
; Sequence 21, Application US/10174364
; Publication No. US20030216308A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2
; CURRENT APPLICATION NUMBER: US/10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1060)
US-10-174-364-21

Query Match 99.7%; Score 1058.8; DB 15; Length 1343;
Best Local Similarity 99.8%; Pred. No. 2,1e-309;
Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGACACACGACGCCCACTGCAAGCAAGCTGCTGCTTGTGTGTCTCCCGGCG 60
DB 2 ATGAGACACACGACGCCCACTGCAAGCAAGCTGCTGCTTGTGTGTCTCCCGGCG 61
QY 61 TCGGCTGCGGCGCTTGGGTTTCGTGCGCGGTGTACTACAGACTCTTGTGCTGCGGT 120

DB 62 TGGGCTGCGGCGCTTGGGTTTCGTGCGCGGTGTACTACAGACTCTTGTGCTGCGGT 121
QY 121 TTACGAGCAATATCTTGAACATGATCATCTCTCCAGCTGATGAGCAAGAGCAGAG 180
DB 122 TTACGAGCAATATCTTGAACATGATCATCTCTCCAGCTGATGAGCAAGAGCAGAG 181
QY 181 TCTCTCTAACAATATCTTGTGACATGCTGCTGCGGACATCTTGGTCTCTTTTTCAT 240
DB 182 TCTCTCTAACAATATCTTGTGACATGCTGCTGCGGACATCTTGGTCTCTTTTTCAT 241
QY 241 GTGTTGTGGAATCTCTGTTGGAAGATTCATCTTGAACATGACAGATGCTGAGTCCC 300
DB 242 GTGTTGTGGAATCTCTGTTGGAAGATTCATCTTGAACATGACAGATGCTGAGTCCC 301
QY 301 GACAAATCATAGAAAGTGTGGAATTCATCTATCAACACTCCATATGATATGATGTA 360
DB 302 GACAAATCATAGAAAGTGTGGAATTCATCTATCAACACTCCATATGATATGATGTA 361
QY 361 CGGTATACATATGACAGATATATCGGTGTGCGCAACCGCTGAAATGACAGCGGTGTA 420
DB 362 CGGTATACATATGACAGATATATCGGTGTGCGCAACCGCTGAAATGACAGCGGTGTA 421
QY 421 TACCCAGCCCGGACCCGGAAGTCAATGTAAGTATTAATCATGACCTGCTTCCGACAG 480
DB 422 TACCCAGCCCGGACCCGGAAGTCAATGTAAGTATTAATCATGACCTGCTTCCGACAG 481
QY 481 ATCCCTATTAATGATGAGTGTGCGCAACATCTGAGTGAAGACTATCATGACCTGTGAT 540
DB 482 ATCCCTATTAATGATGAGTGTGCGCAACATCTGAGTGAAGACTATCATGACCTGTGAT 541
QY 541 CAGCTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 542 CAGCTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
QY 601 ATCTTGAATCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 602 ATCTTGAATCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
QY 661 TACTCCAGGGGAGAGACACCGCATCTTGTTCACCATTAACCTCCATCTTGTGACACTT 720
DB 662 TACTCCAGGGGAGAGACACCGCATCTTGTTCACCATTAACCTCCATCTTGTGACACTT 721
QY 721 TGGGCCCCCGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 722 TGGGCCCCCGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
QY 781 TGGCTGTGACATCATGTGTCGACATGTCGACATGTCGACATGTCGACATGTCGACATC 840
DB 782 TGGCTGTGACATCATGTGTCGACATGTCGACATGTCGACATGTCGACATGTCGACATC 841
QY 841 AACTCTTCTCTACTGCTTCTCATCAGCAAGCGGTTCGCGACATGAGCGCGCAGCTC 900
DB 842 AACTCTTCTCTACTGCTTCTCATCAGCAAGCGGTTCGCGACATGAGCGCGCAGCTC 901
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACCAATCTTAACCTTTTTC 960
DB 902 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACCAATCTTAACCTTTTTC 961
QY 961 ATACAGAGTACCCCTGATCTGCGCGGCAACTCATCTGATCTCAAGATGCTGTATC 1020
DB 962 ATACAGAGTACCCCTGATCTGCGCGGCAACTCATCTGATCTCAAGATGCTGTATC 1021
QY 1021 CAGTATGACAAATAAGAAAACTATATAAGTATCCCGGTGA 1062
DB 1022 CAGTATGACAAATAAGAAAACTATATAAGTATCCCGGTGA 1063

RESULT 9
US-10-246-583-21
; Sequence 21, Application US/10246583
; Publication No. US20040058862A1
; GENERAL INFORMATION:

```

; APPLICANT: Majmuder
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1060)
; US-10-246-583-21

```

```

Query Match      99.7%; Score 1058.8; DB 16; Length 1343;
Best Local Similarity 99.8%; Pred. No. 2.1e-309;
Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ATGAGACACAGCAGCGCCACTCGAGAGCAAGCTCGCTCTTGTGTGCTCCCGGC 60
DB 2 ATGAGACACAGCAGCGCCACTCGAGAGCAAGCTCGCTCTTGTGTGCTCCCGGC 61
QY 61 TCGGCTCGCGCTTGGGTTTCGTGCGCGTGTCTACTACAGCTCTTGTGTGCTCGGT 120
DB 62 TCGGCTCGCGCTTGGGTTTCGTGCGCGTGTCTACTACAGCTCTTGTGTGCTCGGT 121
QY 121 TTACGAGCAAAATCTTGAAGAGATCATCCCTCCAGCTGTGGCAAGAGAGAG 180
DB 122 TTACGAGCAAAATCTTGAAGAGATCATCCCTCCAGCTGTGGCAAGAGAGAG 181
QY 181 TCTCTCAACATATCTTGTGCACTCGCTGCGGACATTTGGTCTCTTTTCTA 240
DB 182 TCTCTCAACATATCTTGTGCACTCGCTGCGGACATTTGGTCTCTTTTCTA 241
QY 241 GTGTTGTGAGCTTCTGTGTGAAGATTTCACTTTGAACATGACAGATGCTCAGGTCCC 300
DB 242 GTGTTGTGAGCTTCTGTGTGAAGATTTCACTTTGAACATGACAGATGCTCAGGTCCC 301
QY 301 GACAAATCATAGAGTCTGAGATTTCTATCCATCCACCTCCATATGATTAAGTGA 360
DB 302 GACAAATCATAGAGTCTGAGATTTCTATCCATCCACCTCCATATGATTAAGTGA 361
QY 361 CCGTTAAACATGACAGGTATATGCTGTGCGACCGGCTCAAGTACACAGGTTCTCA 420
DB 362 CCGTTAAACATGACAGGTATATGCTGTGCGACCGGCTCAAGTACACAGGTTCTCA 421
QY 421 TACCCAGCCGACCCGGAAGAGTCAATGTAGTGTGTTAATCACTCTGTTCCAGCAGC 480
DB 422 TACCCAGCCGACCCGGAAGAGTCAATGTAGTGTGTTAATCACTCTGTTCCAGCAGC 481

```

```

QY 481 ATCCCTATTACTGTTGGCCCAACATCTGAGTGAAGACTATACATGACCTCTGTGAT 540
DB 482 ATCCCTATTACTGTTGGCCCAACATCTGAGTGAAGACTATACATGACCTCTGTGAT 541
QY 541 CAGCTCTCATGTGATTCATGCTTCAACCGTCTACCTGTGAGCGCTGCTCCATCTTCTC 600
DB 542 CAGCTCTCATGTGATTCATGCTTCAACCGTCTACCTGTGAGCGCTGCTCCATCTTCTC 601
QY 601 ATCTTGAATCATATCTTGTGACAAAGTCAAGAGAGAAAGCAATTTTGTCTCGTGGC 660
DB 602 ATCTTGAATCATATCTTGTGACAAAGTCAAGAGAGAAAGCAATTTTGTCTCGTGGC 661
QY 661 TACTCCAGGGGAGAACACACCGCATCTTGTTCACATTAATCTCATCTTGTTCACACTT 720
DB 662 TACTCCAGGGGAGAACACACCGCATCTTGTTCACATTAATCTCATCTTGTTCACACTT 721
QY 721 TGGGCCCCCGCATATCATGATTTCTTTACCACTCTATGGGGGCCCATCCAGAACCGC 780
DB 722 TGGGCCCCCGCATATCATGATTTCTTTACCACTCTATGGGGGCCCATCCAGAACCGC 781
QY 781 TGGCTGTGACATGATGTCCGACATTTGCGCAATGCTAGCCCTTGTGAACAGCCATC 840
DB 782 TGGCTGTGACATGATGTCCGACATTTGCGCAATGCTAGCCCTTGTGAACAGCCATC 841
QY 841 AACTTCTCTCTACTGCTTATCAGCAAGCGGTTCCGCAACATGCGAGCGCCAGCCTC 900
DB 842 AACTTCTCTCTACTGCTTATCAGCAAGCGGTTCCGCAACATGCGAGCGCCAGCCTC 901
QY 901 AAGCTTTCTTCAAGTCCAGAGAACCACTGTATAGTTTCTTACCAATCAATCAATTTTCC 960
DB 902 AAGCTTTCTTCAAGTCCAGAGAACCACTGTATAGTTTCTTACCAATCAATCAATTTTCC 961
QY 961 ATTAACAAGTACCCCTGATCTCGCCGCAAACTCACTGATCAAGATGCTGTGTAC 1020
DB 962 ATTAACAAGTACCCCTGATCTCGCCGCAAACTCACTGATCAAGATGCTGTGTAC 1021
QY 1021 CAGTATGACAAAATGGAATACTATATAAGTATCCCGTGA 1062
DB 1022 CAGTATGACAAAATGGAATACTATATAAGTATCCCGTGA 1063

```

```

RESULT 10
US-10-689-832-21
; Sequence 21, Application US/10689832
; Publication No. US20040121380A1
; GENERAL INFORMATION:
; APPLICANT: Majmuder, Kamud
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729DIV1
; CURRENT APPLICATION NUMBER: US/10/689,832
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78

```

SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-689-833-21

Query Match 99.7%; Score 1058.8; DB 17; Length 1343;
Best Local Similarity 99.8%; Pred. No. 2.1e-309;
Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
1 ATGAGACACAGCAGCCCACTCGAGCCACAGCTGCTGTTGTTGTTCCCGGC 60
2 ATGAGACACAGCAGCCCACTCGAGCCACAGCTGCTGTTGTTGTTCCCGGC 61
61 TCGGCTGCGGCTTGGGTTTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
62 TCGGCTGCGGCTTGGGTTTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
121 TTACGACGAAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGACAGAG 180
122 TTACGACGAAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGACAGAG 181
181 TCTCTCAACAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGACAGAG 240
182 TCTCTCAACAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGACAGAG 241
241 GTGTTTGTGACATCTCTGTTGGAAGATTTCACTTGAACATGACAGTCCG 300
242 GTGTTTGTGACATCTCTGTTGGAAGATTTCACTTGAACATGACAGTCCG 301
301 GACAAATATGAAATGCTGGAATTTCTATTCATCCACACCTCCATGATGAT 360
302 GACAAATATGAAATGCTGGAATTTCTATTCATCCACACCTCCATGATGAT 361
361 CCGTTAACATGACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
362 CCGTTAACATGACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
421 TACCCAGCCGCAACCGGAAAGTCACTGTAAGTTTAACTCACTGCTTCTG 480
422 TACCCAGCCGCAACCGGAAAGTCACTGTAAGTTTAACTCACTGCTTCTG 481
481 ATCCCTATTTATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
482 ATCCCTATTTATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
541 CAGTCTCATCTGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
542 CAGTCTCATCTGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
601 ATCTTGAACTCAATCTGTTGTAACAGCTCAGAGAGAGCAATTTTGTCTC 660
602 ATCTTGAACTCAATCTGTTGTAACAGCTCAGAGAGAGCAATTTTGTCTC 661
661 TACTCAGAGGAGAGAGCAACCGGCAATCTGTTGTAACAGCTCAGAGAG 720
662 TACTCAGAGGAGAGAGCAACCGGCAATCTGTTGTAACAGCTCAGAGAG 721
721 TGGGCCCCCGGCAATCTGTTGTAACAGCTCAGAGAGAGCAATTTTGTCTC 780
722 TGGGCCCCCGGCAATCTGTTGTAACAGCTCAGAGAGAGCAATTTTGTCTC 781
781 TGGGCTGCTGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
782 TGGGCTGCTGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 841
841 AACTTTTCTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
842 AACTTTTCTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
901 AAGGCTTCTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
```

```
DB 902 AAGGCTTCTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961
QY 961 AATACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 962 AATACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021
QY 1021 CAGTATGACAAATATGAAATGCTGGAATTTCTATTCATCCACACCTCC 1062
DB 1022 CAGTATGACAAATATGAAATGCTGGAATTTCTATTCATCCACACCTCC 1063
```

RESULT 11
US-10-450-590-1

```
; Sequence 1, Application US/10450590
; Publication No. US20040076985A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
; FILE REFERENCE: L10316 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/450,590
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/254,923
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: US 60/280,110
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/299,474
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-450-590-1
```

Query Match 99.6%; Score 1057.4; DB 16; Length 1059;
Best Local Similarity 99.9%; Pred. No. 4.8e-309;
Matches 1058; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
1 ATGAGACACAGCAGCCCACTCGAGCCACAGCTGCTGTTGTTGTTCCCGGC 60
2 ATGAGACACAGCAGCCCACTCGAGCCACAGCTGCTGTTGTTGTTCCCGGC 61
61 TCGGCTGCGGCTTGGGTTTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
62 TCGGCTGCGGCTTGGGTTTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
121 TTACGACGAAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGACAGAG 180
122 TTACGACGAAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGACAGAG 181
181 TCTCTCAACAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGACAGAG 240
182 TCTCTCAACAATATCTTGACAGTATCTCTCCAGCTGGTGGCAAGACAGAG 241
241 GTGTTTGTGACATCTCTGTTGGAAGATTTCACTTGAACATGACAGTCCG 300
242 GTGTTTGTGACATCTCTGTTGGAAGATTTCACTTGAACATGACAGTCCG 301
301 GACAAATATGAAATGCTGGAATTTCTATTCATCCACACCTCCATGATGAT 360
302 GACAAATATGAAATGCTGGAATTTCTATTCATCCACACCTCCATGATGAT 361
361 CCGTTAACATGACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
362 CCGTTAACATGACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
421 TACCCAGCCGCAACCGGAAAGTCACTGTAAGTTTAACTCACTGCTTCTG 480
422 TACCCAGCCGCAACCGGAAAGTCACTGTAAGTTTAACTCACTGCTTCTG 481
481 ATCCCTATTTATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
```

Db	1	ATGAGACACAGCAGCCCACTTCGAGCCAAAGACTCGTGTCTTGGTGTGCCCGGAC	60
QY	61	TGGGCTGGGCTTGGGTTTCGAGCCGGTGTCTACTACAGCTCTGTGGCTGGT	12
Db	61	TCGGCTGGGGCTTGGGTTTGGTGCCTGGTCTTACAGCTCTTTCAGTGGCTCGGT	12
QY	121	TTACCGGCAAAATCTGTGACAGTATCATCTCTCCAGCTGGTGGCAGAAAGACAGAG	18
Db	121	TTACAGCAAAATCTGTGACAGTATCATCTCTCCAGCTGGTGGCAGAAAGACAGAG	18
QY	181	TCTCTCAACATATCTCTGGCACTGGCTGGCGGCAATCTTGGTCTCTTTTCAATA	24
Db	181	TCTCTCAACATATCTCTGGCACTGGCTGGCGGCAATCTTGGTCTCTTTTCAATA	24
QY	241	GTGTTTGGGACTTCTGTGGAAAGATTTCATCTTGAACATGACAGATGCTCAGGTCCC	300
Db	241	GTGTTTGGGACTTCTGTGGAAAGATTTCATCTTGAACATGACAGATGCTCAGGTCCC	300
QY	301	GACAGATCATAGAAAGCTGGAAATTCATCCATCCACACTTCATATGATTACTGTA	360
Db	301	GACAAAGATCATAGAAAGCTGGAAATTCATCCATCCACACTTCATATGATTACTGTA	360
QY	361	CCGTTAACCATTTGA CAGGTATTCGTGTGTGCCA CCGGCTCAAGATCA CAGAGGTCTCA	420
Db	361	CCGTTAACCATTTGA CAGGTATTCGTGTGTGCCA CCGGCTCAAGATCA CAGAGGTCTCA	420
QY	421	TACCAAGCCGCGACCCGGGAAAGTCTATTGTAAAGTGTTCATCACTCGTCTCTGACACAG	480
Db	421	TACCAAGCCGCGACCCGGGAAAGTCTATTGTAAAGTGTTCATCACTCGTCTCTGACACAG	480
QY	481	ATCCCTCATTTACTGTGGCCCAACATCTGGA CTGAAGATGAGCACTGTGGCAT	540
Db	481	ATCCCTCATTTACTGTGGCCCAACATCTGGA CTGAAGATGAGCACTGTGGCAT	540
QY	541	CAGGCTCATCTGGAATCACTGCTTCA CCGTTCATCTGGTGGCTGCTCAATCTTCTTC	600
Db	541	CAGGCTCATCTGGAATCACTGCTTCA CCGTTCATCTGGTGGCTGCTCAATCTTCTTC	600

QY	601	ATCTTGAAC	TCATCAT	TGTGT	CAAGCT	CAGAG	GAGAA	GACAA	ATTTT	CGTTC	CGTGGC	666
DB	601	ATCTTGAAC	TCA	TCTGT	GTG	CAAGCT	CAGAG	GAGAA	GACAA	ATTTT	CGTTC	666
QY	661	TACTTCAC	GCGGAA	AGCA	CCGCA	CTTGT	TCA	CATTAC	CTGCAT	CTTTG	CCACACTT	720
DB	661	TACTTCAC	GCGGAA	AGCA	CCGCA	CTTGT	TCA	CATTAC	CTGCAT	CTTTG	CCACACTT	720
QY	721	TGGGCCCC	CGCAT	TCAT	CA	TGA	TCTTT	TACA	CACTT	TA	TGGGGGG	780
DB	721	TGGGCCCC	CGCAT	TCAT	CA	TGA	TCTTT	TACA	CACTT	TA	TGGGGGG	780
QY	781	TGGCTGTG	TGCAT	CA	TGTCC	CAATTC	GCACAA	CA	ATGCT	TAGCCCTT	CTGAA	840
DB	781	TGGCTGTG	TGCAT	CA	TGTCC	CAATTC	GCACAA	CA	ATGCT	TAGCCCTT	CTGAA	840
QY	841	AACCTCTT	CCTCTA	CTG	CTT	CA	TGCA	GAGGG	GTTC	CGCA	CA	900
DB	841	AACCTCTT	CCTCTA	CTG	CTT	CA	TGCA	GAGGG	GTTC	CGCA	CA	900
QY	901	AAGGCTTT	CTTCA	TA	TGCGA	AGCA	CA	CTGT	TAC	AGCTT	TAC	960
DB	901	AAGGCTTT	CTTCA	TA	TGCGA	AGCA	CA	CTGT	TAC	AGCTT	TAC	960
QY	961	ATAACAG	TAGCCCT	GTGAT	TGGCG	GAGAA	CTCA	CA	CTGCAT	CAAGT	CTGTG	1020
DB	961	ATAACAG	TAGCCCT	GTGAT	TGGCG	GAGAA	CTCA	CA	CTGCAT	CAAGT	CTGTG	1020

Qy	1021	CAGTATGACAAAAATGGAAAACTATTAAAGTATCCCGGA	1062
Dd	1021	CAGTATGACAAAAATGGAAAACTATTAAAGTATCCCGTGA	1062

RESULT 13
AS-10-012-140-10

```
Sequence 10, Application US/10012140
Publication No. US20030009017A1
GENERAL INFORMATION:
APPLICANT: Leibny, Kevin R.
APPLICANT: Kuppeler-Libermann, Rosana
APPLICANT: Guckemann, Maria A.
TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 381552004900
CURRENT APPLICATION NUMBER: US/10/012,140
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: 60/246,768
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,772
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,185
PRIOR FILING DATE: 2000-11-15
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 1526
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (139)...(1200)
US-10-012-140-10
```

```
Query Match      98.9%; Score 1050.8; DB 14; Length 1526;
Best Local Similarity 99.3%; Pred. No. 5.8e-307;
Matches 1055; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      1 ATGAGACACACGCGACCCACCTCGACGCCACAGCTCGCTGTCTTGTGGTCCCGGCG 60
DB      139 ATGAGACACACGCGACCCACCTCGACGCCACAGCTCGCTGTCTTGTGGTCCCGGCG 198
QY      61 TCGGCTGGGCGCTTGGGTTTCGTCGCCCGGTGTACTACAGCTCTTGTGGTCCCGGCG 120
DB      199 TCGGCTGGGCGCTTGGGTTTCGTCGCCCGGTGTACTACAGCTCTTGTGGTCCCGGCG 258
QY      121 TTACGACGAAATCTTGAAGATGATCTCTCCAGCTGGTGGGCAAGAGAGAG 180
DB      259 TTACGACGAAATCTTGAAGATGATCTCTCCAGCTGGTGGGCAAGAGAGAG 318
QY      181 TCCTCTACAACTATCTCTTGGCACTCGCTGCGGACATCTTGTCTCTTTTCATA 240
DB      319 TCCTCTACAACTATCTCTTGGCACTCGCTGCGGACATCTTGTCTCTTTTCATA 378
QY      241 GTGTTTGTGACTTCTCTTGTGAAGATTTGATCTTGAACATGACATGCTTCAGTCCG 300
DB      379 GTGTTTGTGACTTCTCTTGTGAAGATTTGATCTTGAACATGACATGCTTCAGTCCG 438
QY      301 GACAAATCATAGAGTGTGAAATTTCTCATTCATCCACAGCTCCATATGATTAAGTGA 360
DB      439 GACAAATCATAGAGTGTGAAATTTCTCATTCATCCACAGCTCCATATGATTAAGTGA 498
QY      361 CCGTTAACATGACAGATATATGCTGTCTGACCCGCTCAAGTACACAGAGTCTCA 420
DB      499 CCGTTAACATGACAGATATATGCTGTCTGACCCGCTCAAGTACACAGAGTCTCA 558
QY      421 TACCCGCGCGGACCCGGAAGTCAATTTGATGTTTAACTCACTGCTCTTGCAGCAGC 480
DB      559 TACCCGCGCGGACCCGGAAGTCAATTTGATGTTTAACTCACTGCTCTTGCAGCAGC 618
QY      481 ATCCCTATTAAGTGTGCGCAACATCTGAGCTGAAGAGTACATGACAGAGCTGTGAT 540
DB      619 ATCCCTATTAAGTGTGCGCGCAACATCTGAGCTGAAGAGTACATGACAGAGCTGTGAT 678
QY      541 CAGGTCTCATCTGATTCACATGCTTACCGCTTACCGTGTGCTGCTCATCTTCTTC 600
DB      679 CAGGTCTCATCTGATTCACATGCTTACCGCTTACCGTGTGCTGCTCATCTTCTTC 728
```

```
QY      601 ATCTGAATCATCATTTGTTGATCAAGCTCAGGAGGAGAGCAATTTGCTCCGNGC 660
DB      739 ATCTGAATCATCATTTGTTGATCAAGCTCAGGAGGAGAGCAATTTGCTCCGNGC 798
QY      661 TACTCCAGGAGGAGAGCAACCGGATCTTGTTCACCATTAACCTCATCTTTGGCACA 720
DB      799 TACTCCAGGAGGAGAGCAACCGGATCTTGTTCACCATTAACCTCATCTTTGGCACA 858
QY      721 TGGGCCCCCGCATCATGATTTCTTACCACTTATGAGGGCGCCCATCCGAACCCG 780
DB      859 TGGGCCCCCGCATCATGATTTCTTACCACTTATGAGGGCGCCCATCCGAACCCG 918
QY      781 TGGCTGTGACATCATGATCCGATTTGCAACATGAGTCCCTTGGAACAGGACATC 840
DB      919 TGGCTGTGACATCATGATCCGATTTGCAACATGAGTCCCTTGGAACAGGACATC 978
QY      841 AACTCTTCTCTACTGCTTTCATACAGCAAGCGGTTCCGACATGCAAGCGGCGACGCTC 900
DB      979 AACTCTTCTCTACTGCTTTCATACAGCAAGCGGTTCCGACATGCAAGCGGCGACGCTC 1038
QY      901 AAGCTTTCTTCAAGTCCGAGAGCAACCTGTACAGTTCTACACCATCATTAATTTC 960
DB      1039 AAGCTTTCTTCAAGTCCGAGAGCAACCTGTACAGTTCTACACCATCATTAATTTC 1098
QY      961 ATACAGTACGCCCCGATCTTCCGCGGAACTCACTGCACTCAAGATGCTGTGTAC 1020
DB      1099 ATACAGTACGCCCCGATCTTCCGCGGAACTCACTGCACTCAAGATGCTGTGTAC 1158
QY      1021 CAGTATGACAAAATGAGAAACCTATAAAGTATCCCGTGA 1062
DB      1159 CAGTATGACAAAATGAGAAACCTATAAAGTATCCCGTGA 1200
```

```
RESULT 14
US-09-813-432-19
Sequence 19, Application US/09813432
Publication No. US20030148485A1
GENERAL INFORMATION:
APPLICANT: Taupier Jr., Raymond J
APPLICANT: Majmuder, Kamud
APPLICANT: Spaderna, Steven K
APPLICANT: Smithson, Glenda
APPLICANT: Mezes, Peter S
APPLICANT: Vernet, Corine A. M.
TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
FILE REFERENCE: 15966-729
CURRENT APPLICATION NUMBER: US/09/813,432
CURRENT FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,836
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/193,843
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 19
```


Oy	961	ATAACAAGTACGCCCTTGATCTTCGGCGGAAACTACACTGCATCAAGATCCTGATGAC	1020
Dd	961	ATAACAAGTAGCCCTTGATCTTCGGCGGAAACTACACTGCATCAAGATCCTGATGAC	1020
Oy	1021	CAGTAGACAAAATGSAAAACCTATTAAAA	1050
Dd	1021	CAGTAGACAAAATGSAAAACCTATTAAAA	1050
 RESULT 15 US-10-174-364-19			
; Sequence 19, Application US/10174364			
; Publication No. US20030216308A1			
; GENERAL INFORMATION:			
APPLICANT: Anderson et al.			
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME			
FILE REFERENCE: 15966-729CIP2			
CURRENT FILING DATE: US/10/174,364			
PRIOR APPLICATION NUMBER: 2002-06-17			
PRIOR FILING DATE: 2000-03-20			
PRIOR APPLICATION NUMBER: 60/190,768			
PRIOR FILING DATE: 2000-03-20			
PRIOR APPLICATION NUMBER: 60/190,972			
PRIOR FILING DATE: 2000-03-22			
PRIOR APPLICATION NUMBER: 60/191,199			
PRIOR FILING DATE: 2000-03-22			
PRIOR APPLICATION NUMBER: 60/191,947			
PRIOR FILING DATE: 2000-03-24			
PRIOR APPLICATION NUMBER: 60/192,665			
PRIOR FILING DATE: 2000-03-28			
PRIOR APPLICATION NUMBER: 60/192,657			
PRIOR FILING DATE: 2000-03-28			
PRIOR APPLICATION NUMBER: 60/192,984			
PRIOR FILING DATE: 2000-03-28			
PRIOR APPLICATION NUMBER: 60/192,664			
PRIOR FILING DATE: 2000-03-28			
PRIOR APPLICATION NUMBER: 60/192,836			
PRIOR FILING DATE: 2000-03-29			
Remaining Prior Application data removed - See File Wrapper or PALM.			
NUMBER OF SEQ ID NOS: 128			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 19			
LENGTH: 1119			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (1)..(1116)			
US-10-174-364-19			
Query Match 98.7%; Score 1048.4; DB 15; Length 1119;			
Best Local Similarity 99.9%; Pred. No. 2,6e-306;			
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Oy	1	ATGAGCACAGCAGCCACCCTCGCAGCGAACAGCTGTCTGTGGTGTCCCCGCGC	60
Dd	1	ATGAGCACAGCAGCCACCCTCGCAGCGAACAGCTGTCTGTGGTGTCCCCGCGC	60
Oy	61	TGCGCTTGGCTTGGGTTTCGTGCCCCTGCTTAACAAGCCCTTGTGTGGCTGGT	120
Dd	61	TGCGCTTGGCTTGGGTTTCGTGCCCCTGCTTAACAAGCCCTTGTGTGGCTGGT	120
Oy	121	TTACGACGAAATATCTTACAGACTATCATCTCTTCCAGCTGTGGCAGAAGACAAG	180
Dd	121	TTACGACGAAATATCTTACAGACTATCATCTCTTCCAGCTGTGGCAGAAGACAAG	180
Oy	181	TCCTCTCAACAATATCTTGGCACTGCGTGTGTCGACACATCTTGGTCCCTTTTGATA	240
Dd	181	TCCTCTCAACAATATCTTGGCACTGCGTGTGTCGACACATCTTGGTCCCTTTTGATA	240
Oy	241	GAGTTTATGACATCTGCTCTTTTCTGCAAGCTTGTGCTGCTTCTTTTCTTTTCTATA	240

Db 241 GTGTTTGTGGACTTCCTGTTGGAAGATTTCATTGAACTGCAAGTGCCTCAGGTCCC 300
QY 301 GACAAATCATAGAAAGTGTGAAATTCATCCATCCACACTTCATATGATTTACTGTA 360
Db 301 GACAAATCATAGAAAGTGTGAAATTCATCCATCCACACTTCATATGATTTACTGTA 360
QY 361 CGGTTAAACCATGACAGGTATATGGCTGTGCACCCGCTGAAAGTACCAACAGGTCTCA 420
Db 361 CGGTTAAACCATGACAGGTATATGGCTGTGCACCCGCTGAAAGTACCAACAGGTCTCA 420
QY 421 TACCCAGCCCGCACCCGGAAGTCATGTGTAAGTGTTCATACACTGTCTTCTGACAGC 480
Db 421 TACCCAGCCCGCACCCGGAAGTCATGTGTAAGTGTTCATACACTGTCTTCTGACAGC 480
QY 481 ATCCCTATTACTGTGGCCCAACATCTGGAAGTGAACATACAGCACTCTGTGCAT 540
Db 481 ATCCCTATTACTGTGGCCCAACATCTGGAAGTGAACATACAGCACTCTGTGCAT 540
QY 541 CAGGTCTGATCTGGATTCACATGCTTCAACCTGCTACCTGTGCTGCATCTTCTTC 600
Db 541 CAGGTCTGATCTGGATTCACATGCTTCAACCTGCTACCTGTGCTGCATCTTCTTC 600
QY 601 ATCTTGAACATCAATTTGTGTA CAAGCTCAGAGGAAGAACAATTTTCGTCTCGTGC 660
Db 601 ATCTTGAACATCAATTTGTGTA CAAGCTCAGAGGAAGAACAATTTTCGTCTCGTGC 660
QY 661 TACTCAGGGGGAAGACCAACCGCATCTTGTTCACCATTAAGTCACTTTTGCACACTT 720
Db 661 TACTCAGGGGGAAGACCAACCGCATCTTGTTCACCATTAAGTCACTTTTGCACACTT 720
QY 721 TGGGCCCCCGGCAATCATGATTTCTTTACACCTCTATGGGGGCGCCATCCAGAACCG 780
Db 721 TGGGCCCCCGGCAATCATGATTTCTTTACACCTCTATGGGGGCGCCATCCAGAACCG 780
QY 781 TGGCTGTGTCATCATGATGTCGCA CATTTGCCAATGTAGCCCTTCTGAAACAGGCCATC 840
Db 781 TGGCTGTGTCATCATGATGTCGCA CATTTGCCAATGTAGCCCTTCTGAAACAGGCCATC 840
QY 841 AACTTCTTCTCTACTGCTTCATGACGAGCGGTTCCGACATGGGAGCGCCAGCCTC 900
Db 841 AACTTCTTCTCTACTGCTTCATGACGAGCGGTTCCGACATGGGAGCGCCAGCCTC 900
QY 901 AAGGCTTTCTTCAAGTGCAGAAAGCAACCTGTACAGTTCTACCAATCATTAATTTTC 960
Db 901 AAGGCTTTCTTCAAGTGCAGAAAGCAACCTGTACAGTTCTACCAATCATTAATTTTC 960
QY 961 ATTAACAAGTAGCCCTGATCTGCGCGCAAACTCACACTGATCAAGAATGCTGTGTAC 1020
Db 961 ATTAACAAGTAGCCCTGATCTGCGCGCAAACTCACACTGATCAAGAATGCTGTGTAC 1020
QY 1021 CAGTATGACAAAATGSAAAAACCTATATAA 1050
Db 1021 CAGTATGACAAAATGSAAAAACCTATATAA 1050

Search completed: December 15, 2004, 02:53:10
Job time : 544.543 secs

This Page Blank (uspto)

atch	44.2%	Score 469.8;	DB 9;	Length 803;
cal Similarity	89.2%;	Pred. No. 4.2e-114;		
518; Conservative	0;	Mismatches 62;	Indels 1;	Gaps 1;
483 CCCCTATTACTGGTGGCCCAACATCTGGAGCTGAAGACTCATTCACAC-CTCTGGCATTC 541				

Db 766 CCGTAAGCTGGGGGGCCCAACCCCTGAGCTTGAATTCATCAGACCGGTCCATGCAATC 707
 Oy 542 ACCTCTCATCTGATTCACCTGCTTACCGGTCTACCTGCTGCTGCTCATCTTCTTCA 601
 Db 706 ATTCCTTATCTGAATTCACCTGCTTCCGTGTAGCGGTACCTGTTCCATTTCTTCA 647
 Oy 602 TCTTGAATCATCTGTTGATCAAGCTAGAGAGAGAGCAATTTTGTCTCGGTGCT 661
 Db 646 TCTGAATCTCATCTGTTGATCAAGCTAGAGAGAGAGCAATTTTGTCTCGGTGCT 587
 Oy 662 ACTTCAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
 Db 586 ACTTCAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 527
 Oy 722 GGGCCCCCGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 781
 Db 526 GGGCCCCCGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 467
 Oy 782 GGGCTGTGACATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 841
 Db 466 GGGCTGTGACATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 407
 Oy 842 ACTTCTCTCTACCTGCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
 Db 406 ACTTCTCTCTACCTGCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 347
 Oy 902 AGGCTTTCTTCAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 961
 Db 346 AGGCTTTCTTCAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 287
 Oy 962 TAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
 Db 286 TAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227
 Oy 1022 AGTATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1062
 Db 226 AGTATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186

RESULT 2
 A2513631/c 362 bp DNA linear GSS 05-OCT-2000
 LOCUS 1M0359L07R Mouse 10kb plasmid UMGCM library Mus musculus genomic
 DEFINITION Clone UMGCM0359L07 R, genomic survey sequence.
 A2513631
 ACCESSION A2513631.1 GI:10694947
 VERSION GSS.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 362)
 Tsam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 CONTACT Contact: Robert B. Weiss
 UNIVERSITY University of Utah Genomes Center
 RM. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 TEL: 801 585 5606
 FAX: 801 585 7177
 EMAIL: ddunham@genetics.utah.edu
 INSERT LENGTH 10000 Std Error: 0.00
 PLATE: 0359 row: 1 column: 07
 SEQ PRIMER: CACACAGGAAACAGCTATGACC
 CLASS: plasmid ends
 High quality sequence stop: 362.

FEATURES
 source

Location/Qualifiers
 1..362
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UMGCM0359L07"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb Plasmid UMGCM library"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrolytically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 24.7%; Score 262; DB 8; Length 362;
 Best Local Similarity 91.4%; Pred. No. 1e-58;
 Matches 277; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
 Oy 760 GGGGGCCCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819
 Db 361 GAGACACCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 302
 Oy 820 GCCCTTGAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879
 Db 301 GCCCTTGAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 242
 Oy 880 ACCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
 Db 241 ACCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 182
 Oy 940 TACACCATCATCTTTCATTAACAAGTACCCCGATCTCGCCGCAAACTCACAC 999
 Db 181 TATACCAACCATTAATCTTTCATTAACAAGTACCCCGATCTCGCCGCAAACTCACAC 122
 Oy 1000 TGCATCAAGATGCTGTGTACCAAGTATGACAAAGAGAGAGAGAGAGAGAGAGAG 1059
 Db 121 TGCATCAAGATGCTGTGTACCAAGTATGACAAAGAGAGAGAGAGAGAGAGAGAG 62
 Oy 1060 TGA 1062
 Db 61 TGA 59

RESULT 3
 A0225693/c 448 bp DNA linear GSS 26-SEP-1998
 LOCUS HS_2009_B2_B08_T7 CIT Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens_genomic clone Plate=2009 Col=16 Row=D, genomic survey
 sequence.
 A0225693
 ACCESSION A0225693.1 GI:3650922
 VERSION GSS.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 448)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2009 row: D column: 16
Class: BAC ends
High quality sequence stop: 448.
Location/Qualifiers
1..448
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2009 Col=16 Row=D"
/sex="male"
/clone_1lb="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"
ORIGIN
Query Match 22.7%; Score 241.4; DB 8; Length 448;
Best Local Similarity 93.7%; Pred. No. 3.5e-53;
Matches 251; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 795 CATGTCGACATTCGCAACATGCTAGCCCTTGTGAACACAGCCATCTTCTCTCTA 854
DB 445 CAGGTGACATTCGCAACATGCTAGCCCTTGTGAACACAGCCATCTTCTCTCTA 386
QY 855 CTGCTTCATCAGCAGCGGTTCCGCAACATGCGCGCCAGCCTCAAGGCTTTCTTCA 914
DB 385 CTGCGTCATCAGCAGCGGTTCCGCAACATGCGCGCCAGCCTCAAGGCTTTCTTCA 326
QY 915 GTGCCGAAGCAACCTGTACAGTTCTACACCAATCAATCTTCCATTAACAAGTACCC 974
DB 325 GTGCCGAAGCAACCTGTACAGTTCTACACCAATCAATCTTCCATTAACAAGTACCC 266
QY 975 CTGGAATCTGCGCGCAACCTGACATGATGATGCTGTGACAGATGACAAAA 1034
DB 265 CTGGAATCTGCGCGCAACCTGACATGATGATGCTGTGACAGATGACAAAA 206
QY 1035 TGGAAAACTATTAAGTATCCCGTGA 1062
DB 205 TGGAAAACTATTAAGTATCCCGTGA 178
RESULT 4
B2169352 813 bp DNA linear GSS 11-OCT-2002
LOCUS B2169352
DEFINITION CH230-259E22, rat CHORI-230 Segment 2 Rattus norvegicus genomic clone
ACCESSION B2169352
VERSION B2169352.1 GI:23810547
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 813)
REFERENCE 1
Zhao,S., Shetty,J., Shatman,S., Teague,G., Geer,K., Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,

TIGR,F., de Jong,P. and Frazer,C.M.
Rat BAC End Sequences from Library CHORI-230 MbOI segment
Unpublished (1999)
Other_GSSs: CH230-259E22.TV
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pje@genomail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ordering information.htm). BAC end
plates: http://www.tigr.org/tdb/bac_end/rat/bac_end_intro.html
plate: 259 row: E column: 22
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..813
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-259E22"
/sex="female"
/cell_type="Brain"
/clone_1lb="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"
ORIGIN
Query Match 20.6%; Score 218.8; DB 8; Length 813;
Best Local Similarity 86.7%; Pred. No. 4.5e-47;
Matches 241; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 125 CAGCAAAATCTTTGACAGATGATCTCTCCAGCTGGTGCAGAAAGACAGATCTCT 184
DB 536 CAGCAAAATCTTTGACAGATGATCTCTCTCACTGTAGCGCAAGAAAGATCTCT 595
QY 185 CCTACAACTATCTCTGGGCACTGCTGCGCGCAACATCTTGCTTTTCTATAGTGT 244
DB 596 CCTACAACTATCTCTGGGCACTGCTGCGCGCAACATCTTGCTTTTCTATAGTGT 655
QY 245 TTGTGACTTCTCTGTTGAAGATTTTCATCTTGAACATGAGATGCTCAGTCCCGACA 304
DB 656 TTGTGACTTCTCTGTTGAAGATTTTCATCTTGAACATGAGATGCTCAGTCCCGACA 715
QY 305 AGATCATAGAGGTGCGAATTCATCATCCATCCACACTTCATATGATTAATGATCGT 364
DB 716 AGATCATAGAGGTGCGAATTCATCATCCATCCACACTTCATATGATTAATGATCGT 775
QY 365 TAACCATGACAGATATTCGCTGCTGCGCACCCGCTC 402
DB 776 TAACCATGATAGATATTCGCTGCTGCGCACCCGCTC 813
RESULT 5
CNS02284 806 bp DNA linear GSS 01-SEP-2000
LOCUS CNS02284
DEFINITION Tetradon nigroviridis genome survey sequence T7 end of clone
227D19 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL178069
VERSION AL178069.1 GI:7816126
KEYWORDS GSS, genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

	REFERENCE	Acanthonomorphae; Acanthopterygii; Percomorpha; Tetraodontiformes;
	AUTHORS	Tetraodonotoidea; Tetraodontidae; Tetraodon.
	TITLE	Roeft Croallius,H., Jailton,O., Dasilva,C., Bouneau,L., Fischer,C,, Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J. Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence <i>Nat. Genet.</i> 25 (2), 235-238 (2000)
Db	JOURNAL MEDLINE PUBMED REFERENCE	20296633 10835645
OY	AUTHORS	2 Roeft Croallius,H., Jailton,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis <i>Freshwater Res.</i> 10 (7), 939-949 (2000)
Db	JOURNAL MEDLINE PUBMED REFERENCE	20259837 10899143 3 (bases 1 to 806) Genoscope, Direct Submission
OY	AUTHORS	Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE [E-mail : sequef@genoscope.cns.fr]
Db	JOURNAL TITLE	- Web : www.genoscope.cns.fr This sequenece is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon. location/Qualifiers
FEATURES	SOURCE	1..806 /organism="Tetraodon nigroviridis" /mol_type="genomic DNA" /db_xref="taxon:99883" /clone="227D19" /clone_1fb="G"
ORIGIN		/note="Genoscope sequence ID : CGAG227CBL0LP1-end : T7"
	Query Match	19.8%; Score 209.8; DB 9; Length 806; Best Local Similarity 56.3%; Pred.No. 1.1e-44;
	Matches 418; Conservative	7; Mismatches 289; Indels 29; Gaps 1;
OY		166 GCACGAAGAACAAGAATCTCCCTCAACAATATTCTTTGGACTCGTCGCCGCACATCTTG 225
Db		2 GGAAAGACCAGAGAGCCCCGTACTACTAAGCTCGGCGGAGACAGGCTCCGACATCTTC 61
OY		226 GTCTCTTTTTCATAGTAGTTMTTGGACTTCTGTGGAAATTTCACTTGAATCATGAC 285
Db		62 TTCACAGCTCTTCATCATATCTTGTGGGGCTTCTGTGGAGACGGGGTTTTCCACCGGGAG 121
OY		286 ATGCTCACAGTCCCACAGAACATGATAGTAAGTGTGTGAATTTCATATCATCCACACTTC 345
Db		122 GTCCCGGGGTCTCTTAACTCACTAGTACGAGCCGCGCGAGTTTCCGCCAACCGCTTC 181
OY		346 ATAATGATTAATACTGAACGTBAACATTGACAGGATATATCGTGTCCGACCGGCTCAG 405
Db		182 ATTGTGTCCAACGGTCCCCCTCACCGTGAACGGCTACGTGACGCTGTGCACCCCTTCTC 241
OY		406 TACCAACAGGCTCATACCCAGGCGCCGACCCGGAAGTCAITGTAAAGTGTATTAATCAC 465
Db		242 CACAGGAGNACAGTACCCGCGCCGAGACCGAGAGATCATGCGGTGTCTCTGTGTCTG 301
OY		466 TGCTTCTTGAACAGATCCCTTATTAATGTGTGCCCCAATCATGTGACGTGAAGATCAATC 525
Db		302 TCAGTGGCCCTGCGGCTCCCTTCTTCTGTGTGTGCGACATGTGAAGAGAACGCCACCG 361
OY		526 AGACCTCTGTGATCAACGTCCTCATCTGAATTCATCTGTTAACCGTCTACCTGTGCC 585
Db		362 CCAGAGCGGTGAGACCGCTCTCTCATCTGAACCCAMGTGACCATCATCACTACTCTGCCC 421
OY		586 TGTCAATCTTCTTCACTTGAATCAAATCATTTGTATCAAGCTCAGGAGGAAGAGCAT 645

Db 422 TCACAGCATCTTCTGTCCTCAACTTTGATCATCCAMAAGCTGAGGCTCGGCAGARG 481
 QY 646 TTTGCTCCGTGCTACTCCAC-----GGGAGAA 676
 Db 482 CGGCGAGCGCAGCGGGGGGAGCGGGGCMCAAGTMAACGCTTCGCGCCGCTGGGAGAA 541
 QY 677 CCACCGGCATCTGTTACCATTAACCTCATCTTTGGCACACTTTGGGCCCCCGGCATCA 746
 Db 542 GCACGGCCATGCTGCTGGCCATCACCCTCGTCTTCTGTGCTCTGGGCAACCGAGACG 601
 QY 737 TCATGATCTTTACCACTCTATAGGGGGGCCATCCAGAACCGTGCTGTGCACATCA 796
 Db 602 TGGTGTGATCATACCACTGTACGTGTATGTGGTTTCAACGGAAATGGCGGCTCCACTCG 661
 QY 797 TGTCCGACATTCGCAACATGCTAGGCCCTTGTGAACACAGCCATCAATTCTTCCTACT 856
 Db 662 CCTACGACCTGTCCAAACATGCTGGCCATGCTCAACAGGCGGTAAACTTCTTCCTACT 721
 QY 857 GCTTCATCAAGCAAGGGTTCGCG 879
 Db 722 GCTTCGTCAACAGCGTTCGCG 744

LOCUS	AQ344035	588 bp	DNA	linear	SSS 07-MAY-1995
DEFINITION	RPcII1-124023.TV RPcI-11 Homo sapiens genomic clone RPcI-11-124023, genomic survey sequence.				
ACCESSION	AQ344035				
VERSION	AQ344035.1	GI:4168931			
KEYWORDS	SSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 588) Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.				
TITLE	Use of BAC End Sequences from Library RPcI-11 for Sequence-Ready Map Building				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are derived from the human BAC library RPcI-11. For BAC library availability, please contact Plier de Jong (plier@eig.org, med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html Seq primer: T7 Class: BAC ends.				
FEATURES	Location/Qualifiers				
source	1..588 /morganism="Homo sapiens" /mol_type="genomic DNA" /db_xref="GDB:7547470" /db_xref="taxon:9606" /clone="RPcI-11-124023" /sex="Male" /cell_type="Lymphocytes" /clone_1fb="RPcI-11" /notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPcII1 Human Male BAC library"				
ORIGIN					
Query Match	18.8%;	Score 200;	DB 8;	Length 588;	
Best Local Similarity	100.0%;	Pred. No. 4,3e-42;			

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 CAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGCAGAAGTCTT 184
 DB 201 CAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGCAGAAGTCTT 142

QY 185 CCTACAACTATCTCTTGAGCACTGCTGCTGCCAGACATCTTGGTCTCTTTTTCATAGTGT 244
 DB 141 CCTACAACTATCTCTTGAGCACTGCTGCTGCCAGACATCTTGGTCTCTTTTTCATAGTGT 82

QY 245 TTGTGACCTTCCTGTGGAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCGACA 304
 DB 81 TTGTGACCTTCCTGTGGAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCGACA 22

QY 305 AGATCATAGAGTGTCTGAA 324
 DB 21 AGATCATAGAGTGTCTGAA 2

RESULT 7
 AQ351433/c 603 bp DNA linear GSS 07-MAY-1999
 LOCUS RPI11-113113.TV RPI11-113113 Homo sapiens genomic clone RPI11-113113,
 DEFINITION genomic survey sequence.
 ACCESSION AQ351433
 VERSION AQ351433.1 GI:4178768
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 603)
 AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
 Venter,J.C.
 TITLE Use of BAC End Sequences from Library RPI11 for Sequence-Ready
 Map Building
 JOURNAL Unpublished (1997)
 COMMENT Other_GSSs: RPI11-113113.TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@ligr.org
 Clones are derived from the human BAC library RPI11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genetics (info@resgen.com). BAC end search page:
 http://www.ligr.org/tadb/humgen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source location/Qualifiers
 1..603
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RPI11-113113"
 /sex="Male"
 /cell_type="lymphocytes"
 /clone_lib="RPI11"
 /note="Vector: PBACe3.6; Site 1: ECORI; Site 2: ECORI;
 RPI11 Human Male BAC Library"

ORIGIN
 Query Match 18.5%; Score 197; DB 8; Length 603;
 Best Local Similarity 100.0%; Pred. No. 2.7e-41;
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 125 CAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGCAGAAGTCTT 184

DB 197 CAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGCAGAAGTCTT 138
 QY 185 CCTACAACTATCTCTTGAGCACTGCTGCTGCCAGACATCTTGGTCTCTTTTTCATAGTGT 244
 DB 137 CCTACAACTATCTCTTGAGCACTGCTGCTGCCAGACATCTTGGTCTCTTTTTCATAGTGT 78

QY 245 TTGTGACCTTCCTGTGGAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCGACA 304
 DB 77 TTGTGACCTTCCTGTGGAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCGACA 18

QY 305 AGATCATAGAGTGTCTG 321
 DB 17 AGATCATAGAGTGTCTG 1

RESULT 8
 CNS02BVQ 947 bp DNA linear GSS 01-SEP-2000
 LOCUS Tetradon nigroviridis genome survey sequence pUC-ori end of clone
 DEFINITION 254A20 of library G from Tetradon nigroviridis, genomic survey
 sequence.
 ACCESSION AL190367
 VERSION AL190367.1 GI:7828471
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetradon nigroviridis
 ORGANISM Tetradon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorphia; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.
 REFERENCE 1
 AUTHORS Roest Crolius,H., Jaillon,O., Da Silva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 TITLE Estimate of human gene number provided by genome-wide analysis
 using Tetradon nigroviridis DNA sequence
 JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
 MEDLINE 20296633
 PUBMED 10835645

REFERENCE
 AUTHORS Roest Crolius,H., Jaillon,O., Da Silva,C., Ozouf-Costaz,C.,
 Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
 Saurin,W., Bernot,A. and Weissenbach,J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetradon nigroviridis
 JOURNAL Genome Res. 10 (7), 939-949 (2000)
 MEDLINE 20359837
 PUBMED 10899143

REFERENCE
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetradon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetradon.

FEATURES
 source location/Qualifiers
 1..947
 /organism="Tetradon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="254A20"
 /clone_lib="G"
 /note="Genoscope sequence ID : C0AG254BA10SP1-end ;
 pUC-ori"

ORIGIN
 Query Match 17.7%; Score 187.6; DB 9; Length 947;
 Best Local Similarity 70.0%; Pred. No. 1e-38;
 Matches 266; Conservative 3; Mismatches 102; Indels 9; Gaps 1;
 125 CAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGCAGAAGTCTT 184

```

QY 692 TCACCATTAAGTCTGATCTTGGACACTTGGGCCCCCGCATCAATGATTTTACC 751
DB 7 TGGCATTAAGTCTGATCTTGGACACTTGGGCCCCCGCATCAATGATTTTACC 66
QY 752 ACCCTATGAGGGGCGCCATTCAGAACCGCT-----GGCTGATGACATCATGTCG 802
DB 67 ACTTTACTGCGCTCGCGCAGCCTGCGGGGCGCGCGCGGTGTGCACTTCTCAACG 126
QY 803 ACATGACCAAGATGCTAGCGCTTGGACACAGCATCACTTCTTCTTACTGCTTCA 862
DB 127 ATCTTGGCAACATGTTGGCGTGTGCTCAACAGCGGTGTTACTTCTTCTTACTGCTTCA 186
QY 863 TCAGCAAGCGGTTCCGACCATGAGCGCGCGCATCGCTTCTTCAAGTCCAGA 922
DB 187 TCAGCAAGCGGTTCCGCGGCGCATGCGCGCAACGTGCTCGAGCGCTGCTCACTGCAAGA 246
QY 923 AGCAACCTGATACGTTTCTTACACCAATCAATCTTTCCATTAAGAGCCCTGATCT 982
DB 247 AGCAGCTGCAAGCGCTTCTTACCGCAGCGCAACTTTCATCAAGAGTGCCTGCTG 306
QY 983 CGCGGCAAACTCAACATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1042
DB 307 CACGAGAACTCCCACTGATCAAGATGATGATGATGATGATGATGATGATGATGATGAT 366
QY 1043 CTATTAAGATATCCCGCTGA 1062
DB 367 CCATCTGATATTCCTCTGA 386

RESULT 9 708 bp mRNA linear EST 20-FEB-2003
BU610327
LOCUS BU610327
DEFINITION U1-M-DJ2-bw1-f-07-0-UT.r1 NIH BWAP DJ2 Mus musculus cDNA clone
ACCESSION BU610327
VERSION BU610327.1 GI:23276542
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 708)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
JOURNAL MEDLINE PUBMED
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestr@mail.nih.gov
Tissue Procurement: Dr. Robin Davison
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.regen.com).
Seq primer: M13 REVERSE.
Location/Qualifiers
1..708
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="U1-M-DJ2-bw1-f-07-0-UT"
/tissue_type="neofornical organ and postrema"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="NIH_BWAP_DJ2"

```

```

/Note="Organ: Brain; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
U1-M-DJ2 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is: GCTCATCATGAT, subformical
organ and area postrema."

ORIGIN
Query Match 17.4%; Score 185; DB 5; Length 708;
Best Local Similarity 92.8%; Pred.No. 4,6e-38;
Matches 194; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 854 ACTGCTTATCAGCAAGCGGTTCCGACCATGAGCGCGCATGAGGCTTTCTTCA 913
DB 1 ACTGCTTATCAGCAAGCGGTTCCGACCATGAGCGCGCATGAGGCTTTCTTCA 60
QY 914 AGTGCAGAAAGCAACCTGATCAAGTTCTACCAATCATTAATTTTCATTAAGATGCC 973
DB 61 AGTGCAGAAAGCAACCTGATCAAGTTCTACCAATCATTAATTTTCATTAAGATGCC 120
QY 974 CCGATGATCGCGGCAAACTCAACATGATGATGATGATGATGATGATGATGATGATGAT 1033
DB 121 CCGATGATCGCGGCAAACTCAACATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 1034 ATGAAAACCTATTAAGATATCCCGCTGA 1062
DB 181 ATGAAAACCTATTAAGATATCCCGCTGA 209

RESULT 10 824 bp DNA linear GSS 05-JAN-2004
CU137737
LOCUS CU137737
DEFINITION ISB1-110A2.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-110A2,
genomic survey sequence.
ACCESSION CU137737
VERSION CU137737.1 GI:4061372
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 824)
Kremnitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: rkwilson@wustl.edu
Insert Length: 7500 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGG
Clase: BAC ends
High quality sequence start: 2
High quality sequence stop: 672.
Location/Qualifiers
1..824
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-110A2"
/clone_lib="ISB1"
/Note="Vector: pBel0BAC11; ISB-1 Xenopus tropicalis BAC
library segment 1"

```

ORIGIN

Query Match 17.2%; Score 183; DB 9; Length 824;
Best Local Similarity 59.5%; Pred. No. 1.7e-37;
Matches 309; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

125 CAGCAATATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGACAGAAAGCTT 184
167 CAGTGAACATCTGACTGCTGTGGACACTTCAAGGCTGGCATGAGAACAAAGATCTT 226
185 CCTACAACTATCTCTTGGCACTGCTGCTCCGACATCTTGTCTCTTTTTCATAGTGT 244
227 CTTACAGTTACTCTGTGGCTCTTACCATCTCAGACATCTTGGACGAGATTTTCATATT 286
245 TTGTGACCTTCTGTTGAAAGATTTTATCTTGAACATGACGATGCTCAGTCCCGACA 304
287 TTGTGGCTTTTATCTGCAACAGCAATCTCCACCGCAAGGTGCCAGTACCTTATCC 346
305 AGATCATAGAAAGCTGGAATTTCTCATCTCCATCCACCTCATATGATTTACTGATCC 364
347 ATGTGTCAGTGTCTTGAAGTCTCTCTTATATGATCATTTATTTGGTGAACATGATAC 406
365 TAAACATTTGACAGTATATGCTGTGCTCCACCGCTCAAGTACCAACGCTCTCAATCC 424
407 TGACCGTGAACCGTTATGTGGCATTTATGCTATCCGATACAAATTCGCTTTCTTACC 466
425 CAGCCGCAACCCGGAAGTATTTGAAGTGTTCATCATCTGCTTCTGACAGCATCC 484
467 CAAAGGCAACCCGTAGAGTATTTGTCTTCTCTCATGTTTGTGACCGGTGTAC 526
485 CCTATTACTGTGTGGCCCAACATCTGAGCTGAGCTACATCAGACACTGTGTGATCAG 544
527 CTTTCTACTGTGTGAGTATGTTTGAAGGATTCGCGTACCCAGATTTCTGACCTTA 586
545 TCTCTACTGTGATTCACATGCTTCACTGATGCTGCTGCTCTGCTTCTTCACTT 604
587 TACTCAAGTGAACCACTGCTTTATATATATCTTCACTTCCCTGACCATATTTTGTAT 646
QY 605 TGAATCATCATCTTGTGTACAGCTCAGAGGAAAGACA 643
DB 647 CCAACTCTGTCTATATCTACAGGCTGAGAAAAAGACA 685

RESULT 11
LOCUS CL649540 1269 bp DNA linear GSS 06-JUL-2004
DEFINITION CH213-237F13.SP6 CH213 Gastrosteus aculeatus genomic clone
ACCESSION CL649540
VERSION CL649540.1 GI:49668964
KEYWORDS GSS.
SOURCE Gastrosteus aculeatus (three spined stickleback)
ORGANISM Gastrosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gastrosteiformes;
Gastrosteidae; Gastrosteus.
1 (bases 1 to 1269)
REFERENCE Kingsley, D., Grimwood, J., Dickson, M., Schmutz, J. and Myers, R.M.
Expressed sequence tags from Gastrosteus aculeatus
Unpublished (2004)
JOURNAL COMMENT Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Avenue, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@hgc.stanford.edu
Plate: 237
Class: BAC ends
High quality sequence start: 19
High quality sequence stop: 869.
Location/Qualifiers 1..1269

FEATURES
source

/organism="Gastrosteus aculeatus"
/mol_type="genomic DNA"
/strain="Salmon River"
/db_xref="taxon:69293"
/clone="CH213-237F13"
/sex="Mixed"
/cell_type="Blood"
/clone_1lb="CH213"
/note="Vector: pTARBAC2.1; Site_1: EcoRI. The sequence of the clone was established as a mapping and sequencing collaboration at the Stanford Genome Evolution Center, funded by the NIH Center of Excellence in Genomic Science (CEGS) initiative (<http://cegs.stanford.edu>). The clone was isolated from the BAC library CHORI-213 built by Pletzer deLong in collaboration with the Stanford Genome Evolution Center (<http://www.chori.org/bacpac/>). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm)."

Query Match 16.2%; Score 172.4; DB 9; Length 1269;
Best Local Similarity 75.2%; Pred. No. 1.3e-34;
Matches 215; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

777 CCGCTGCTGTGACATCATGTCCGACATTCGCAACAGCTTCTGAAACACAGC 836
42 CGGGCAGCTCTCCAGCTCTCAACGACCTGACATTATGCTGCTCTCAACACCG 101
837 CATCAACTTCTCTTACTGCTTCTACAGCAAGCGTTCCGACCATGGACCGCCAC 896
102 GGTCAACTTCTTCTTACTGCTTCTACAGCAAGGTTTCCGGGATGGCGGCAACGT 161
897 GCTCAAGCTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACCAATCATATT 956
162 CCGCGGGCCCTGTGTAACCTGCGGAGGACGCGCCGTTCTACGCGACCAACTT 221
QY 957 TTCATTAAGAGCCCTGATCTGCGGCAACCTCACTGATCAAGTGTGT 1016
DB 222 TTCATCAACAGAGCCGCTGATCTCACCGGCAACTCCACTGATTAAGTGTGT 281
QY 1017 GTACAGTATGACAAAATGGAACCTTAAAGTATCCCGTCA 1062
DB 282 GTACAGTATGACAAAATGGAACCCGCTGTATTTCTCTTGA 327

RESULT 12
LOCUS CL137848 867 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-110C15.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-110C15,
genomic survey sequence.
ACCESSION CL137848
VERSION CL137848.1 GI:40631483
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 867)
REFERENCE Krenitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
JOURNAL COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@atson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATCAGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 11
High quality sequence stop: 673.
Location/Qualifiers

FEATURES

SOURCE

1..867
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB-110C15"
/note="Vector: pBelOBAC11, ISB-1 Xenopus tropicalis BAC library segment 1"

ORIGIN

Query Match 14.1%; Score 150; DB 9; Length 867;
Best Local Similarity 58.1%; Pred. No. 1,1e-28;
Matches 264; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 125 CAGCAAAATCTTGAAGATGATCATCTCTCCACTGTGGCAAGAAACAAGATCTT 184
DB 182 CAGTGAACATGCTGACTGCTGTGGCACTCTCAAGCTGCAATCGAAGCAAGAAATCTT 241
QY 185 CCTACACATATCTCTTGGCACTGCTGTGGCAATCTTGGATCTCTTTTCATAGTCT 244
DB 242 CCTACAGTTACTCTGTGGCTCTTACATCTGAGACATTTGAGCAATTTTCATCTT 301
QY 245 TTGTGGACTCTCTGTGGAAATTCATCTTGAACATGACAGATCTCTAGGTCGCCGACA 304
DB 302 TTGTGGGCTTTATCTCTGCAAAAGCAATATCTCCAGAGTCCCAAGTACCTTAATCC 361
QY 305 AGATCATGAAAGTGTGGAATTTCTATCATCATCCACCTCCATATGATTAATGATCCGT 364
DB 362 ATGTGGTCAAGTCTTGAATTTCTCTTATATCATGATCTATTTGGGTACAGATGATAC 421
QY 365 TAACCATGACAGGAATATCGCTGTGCTGCCACCGCTCAAGTACACAGAGTCTCATACC 424
DB 422 TGACCGTGAACCGTTATGTGGCAATATGCTATCGCTACAAATATCTCTCTCTTAC 481
QY 425 CAGCCCGACCCGGAAGATCATTTGAATGTTATCATCATCTCTCTGCAACAGATC 484
DB 482 CAGAGCAACCCGTAAGATGATCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 541
QY 485 CCTATTAGTGTGGCCCAATCATCTGATGATGATGATGATGATGATGATGATGATGATG 541
DB 542 CCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 544
QY 545 TCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601
DB 602 TACTCAGTGAACCACTGCTTTATATATATCTT 635

RESULT 13
CNS02ENP 898 bp DNA 11linear GSS 01-SEP-2000
LOCUS
DEFINITION Tetradodon nigroviridis genome survey sequence PUC-ori end of clone
sequence.
262P08 of library G from Tetradodon nigroviridis, genomic survey
ALI93966
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Tetradodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Acanthomorpha; Acanthopterygii; Teleostei; Euteleostei; Neoteleostei;
Tetraodontidae; Tetraodontiformes; Tetradodon.

REFERENCE
AUTHORS
Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fzames, C., Wincker, P., Brothier, P., Quetier, F.,
Sautin, W., and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetradodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
JOURNAL
MEDLINE
PUBMED
REFERENCES
AUTHORS
Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C.,
2
10835645

TITLE
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Sautin, W., Bernot, A., and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
JOURNAL
MEDLINE
PUBMED
REFERENCES
AUTHORS
3 (bases 1 to 898)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetradodon.
Location/Qualifiers
1..898
/organism="Tetradodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="262F08"
/clone_11b="G"
/note="Genoscope sequence ID : C0AG262DC048P1-end :
PUC-ori"

COMMENT

FEATURES

ORIGIN

Query Match 14.0%; Score 149; DB 9; Length 898;
Best Local Similarity 56.9%; Pred. No. 2e-28;
Matches 292; Conservative 0; Mismatches 220; Indels 1; Gaps 1;

QY 121 TTACAGCAATATCTTGAAGATGATCATCTCTCCAGTGTGGCAAGAAACAAG 180
DB 370 TTCCAGTGAACATCTGACGGGGTGGCTTACCCGCTGCGCAACAAAG 429
QY 181 TCTCTCAACATCTCTTGGCACTGCTGCTGCGCAATCTTGTCTCTTTTCTATA 240
DB 430 GCACTGATCACTACCTGCGGGTGAACAGGCTCCGATCTCTCCAGCTCTTCAAC 489
QY 241 GTGTGTGACTCTCTGTGGAAATTCATCTTGAATGATGATGATGATGATGATGATG 300
DB 490 ATCTGTGGGCTTCTGTGGAGAGGCAATTTCCACGGGAGTCCCGCTCTCTC 549
QY 301 GACAATATATGAAAGTGTGAATTTCTATCATCATCCACCTCCATATGATTAAGTA 360
DB 550 TTACACTCACTGACGCCCGCGAGTTGGCGGCAACAGGCTCATCTGTCCACGCTC 609
QY 361 CCGTTACCATTTGACAGATATTCGCTGTGCGCCACCGCTTCAATACACAGGTTCA 420
DB 610 CCCCTACCGTGAACCGCTGACGTGGCTGTGCGCACCCCTCTTCCACAGGCAATCAGC 669
QY 421 TACCCAGCCCGACCCGGAAGTCAATTTGAATGTTTATCATCATCTGCTTCTGACAGC 480
DB 670 TACCCGCGCCGAGACAGAGATCATCGCGGTGGTCTGTGTGCTGTGCTGCTGCGGC 729
QY 481 ATCCCTATTACTGTGGCCCAAGATCTGACCTGAAGATCAATCAAGCACTTGTGCAT 540
DB 730 CTGCGCTTCTTGTGTGTGCGAATGAGAGAAACAGCCAC-CCGCGACGGGCTGAGA 788
QY 541 CAGCTCTCATCTGATGATCACTGCTTCACTCTTCACTGATGATGATGATGATGATG 600
DB 789 CCGCTTCTCATCTGAGACCAAGTCAATCTTCACTTCTGCTGCTGCTGCTGCTGCTG 848
QY 601 ATCTGAATCAATCTTGTGTATCAAGTCAAG 633
DB 849 GTCTCACTCTTGTATCATCAACAGTCAAG 881

RESULT 14
AL921815 515 bp mRNA 11linear EST 06-JUL-2004
LOCUS
DEFINITION AL921815 PUR-Z142 Danio rerio cDNA clone 101-F03-2, mRNA sequence.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 18:35:19 / Search time 56.6035 Seconds
(without alignments)
2237.166 Million cell updates/sec

Title: US-09-995-225B-16

Perfect score: 1872
Sequence: 1 MENTHAHLAANSLSLWSMSPG.....CIKMLVYQYDKNGKPKVSP 353

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358722929 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23sep04:*
1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003ba:*
8: geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1872	100.0	353	5	AU76416 Novel G-P
2	1872	100.0	353	5	ABU04075 Human G-P
3	1872	100.0	353	5	ABB79517 Human che
4	1872	100.0	353	6	ABP98724 Human AXO
5	1872	100.0	353	6	ABR44437 Human G-P
6	1872	100.0	353	6	ABB82499 Human TGR
7	1872	100.0	591	7	ADP70485 Orphan re
8	1868	99.8	353	4	AU10068 Chemokine
9	1868	99.8	353	7	ADU87769 G-coupled
10	1868	99.8	353	8	AD179325 NOVI1 pro
11	1868	99.8	353	8	AD056004 Human NOV
12	1868	99.8	353	8	AD028954 Human nov
13	1857	99.2	372	4	AU10067 Chemokine
14	1857	99.2	372	5	AAE18645 Human G-P
15	1857	99.2	372	6	ABR62521 Human G-P
16	1857	99.2	372	7	ADD18023 Human G-P
17	1857	99.2	372	7	ADJ87767 G-coupled
18	1857	99.2	372	8	AD179323 NOVI0a pr
19	1852	98.5	372	8	AD056002 Human NOV
20	1844	98.5	353	5	AAE29236 Human NOV
21	1844	98.5	353	6	ABG71163 Human hum
22	1841	98.3	369	6	ABU09571 Human pro
23	1841	98.3	369	6	ABR62522 Human G-P
24	1756	93.8	345	6	ABB82508 Mouse TGR
25	1756	93.8	345	8	AD028956 Mouse nov

ALIGNMENTS

RESULT 1	
ID AU76416 standard; protein; 353 AA.	
XX	
AC AU76416;	
XX	
DT 08-MAY-2002 (first entry)	
XX	
DE Novel G-protein coupled receptor TGR8.	
XX	
KW G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;	
XX	
KW cytostatic; antiinflammatory; anticancer; fetal brain;	
XX	
KW central nervous system disease; circulatory organ disorder; cancer;	
XX	
KW metabolic disease; immunological disease; gastrointestinal disease;	
XX	
KW gene therapy; transgenic animal; human.	
XX	
OS Homo sapiens.	
XX	
PN MO200194582-A1.	
XX	
PD 13-DEC-2001.	
XX	
PF 01-JUN-2001; 2001WO-JP004643.	
XX	
PR 02-JUN-2000; 2000JP-00170446.	
XX	
PR 23-JUN-2000; 2000JP-00194928.	
XX	
PA (TAKEDA) TAKEDA CHEM IND LTD.	
XX	
PI Terao Y, Matsui H, Shintani Y;	
XX	
DR WPI; 2002-164317/21.	
XX	
DR N-PSDB; ABR15562, ABR15563.	
XX	
PT Human fetal brain-originated G-protein-coupled receptor protein TGR8 and	
XX	
PT encoding DNA, for developing drugs to treat e.g. diseases of the central	
XX	
PT nervous system or circulatory organs, cancer, and metabolic diseases.	
XX	
PS Claim 1, Fig 2; 102pp; Japanese.	
XX	
CC The invention describes a human fetal brain-originated G-protein-coupled	
XX	
CC receptor protein, or its salt. The protein and encoded DNA are useful for	
XX	
CC developing drugs to treat e.g. diseases of the central nervous system or	
XX	
CC circulatory organs, cancer, metabolic diseases, immunological diseases	
XX	
CC and gastrointestinal diseases. The invention also describes creation of a	
XX	
CC probe for gene therapy and construction of a transgenic animal. This is	
XX	
CC the amino acid sequence of the novel G-protein coupled receptor TGR8,	
XX	
CC described in the method of the invention	

26	1643	87.8	318	7	ADJ87772	Adj87772 G-coupled
27	1643	87.8	318	8	AD179348	Adi79348 NOVI0b pr
28	1643	87.8	318	8	AD056027	Ado56027 Human NOV
29	1638	87.5	333	5	AAE17081	Aae17081 Human G-P
30	1638	87.5	333	7	ABW00814	Abw00814 Human GPC
31	1638	87.5	343	5	ABB79519	Abb79519 Human che
32	1638	87.5	356	5	ABB79518	Abb79518 Human che
33	1638	87.5	385	5	ABP95606	Abp95606 Human GPC
34	1626	86.9	333	5	AAE17229	Aae17229 Human tly
35	1471	78.6	287	4	AAU25559	Aau25559 Human G-P
36	1421	75.9	321	7	ADC12696	Adc12696 Human GPC
37	1317	70.4	265	6	ABR81706	Abp81706 Human G-P
38	1317	70.4	265	6	ABR62524	Abp62524 Human G-P
39	912.5	48.7	313	4	AAU25556	Aau25556 Human G-P
40	742	39.6	340	8	AD028936	Ado28936 Mouse nov
41	742	39.6	348	6	ABR82431	Abp82431 Murine 18
42	737	39.4	374	5	ABU04070	Abu04070 Human G-P
43	737	39.4	374	5	AAU49155	Aau49155 Human G-P
44	737	39.4	374	6	ABR44438	Abp44438 Human G-P
45	737	39.4	374	6	ABB82505	Abb82505 Human TGR

```

XX SQ Sequence 353 AA;
Query Match 100.0%; Score 1872; DB 5; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.8e-200;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSLSWSPSGACGFGVPPVYYSLLCLGLPANILTVIILSQLVARRQK 60
DB 1 MEHTAHLAANSLSWSPSGACGFGVPPVYYSLLCLGLPANILTVIILSQLVARRQK 60
QY 61 SSYNYLLAALAAADILVLFVIVDFLEDFILNQMPQVPDKIEVLEFSSIHHSIMWTV 120
DB 61 SSYNYLLAALAAADILVLFVIVDFLEDFILNQMPQVPDKIEVLEFSSIHHSIMWTV 120
QY 121 PLTIDRYAVCHPLKHTVSPARTKVIYSVYTTCPLTSPYWMWPNWTEYISTSVH 180
DB 121 PLTIDRYAVCHPLKHTVSPARTKVIYSVYTTCPLTSPYWMWPNWTEYISTSVH 180
QY 181 HVLWICFTVYLVPCSIFFILNSIIYKLRKSNPFLRGYSTGKTALILFTTSTPATL 240
DB 181 HVLWICFTVYLVPCSIFFILNSIIYKLRKSNPFLRGYSTGKTALILFTTSTPATL 240
QY 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFRMAAATL 300
DB 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFRMAAATL 300
QY 301 KAFKCKQKOPVQFTYNNHFSITSSPWSPANSHCIKMLVYQYDNGKRPKVSP 353
DB 301 KAFKCKQKOPVQFTYNNHFSITSSPWSPANSHCIKMLVYQYDNGKRPKVSP 353

RESULT 2
ABJ04075 ID ABJ04075 standard; protein; 353 AA.
XX AC ABJ04075;
XX DT 11-OCT-2002 (first entry)
XX DE Human G protein coupled receptor hRUP35.
XX KM Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
XX KW hRUP2; hRUP3; hRUP34; hRUP35; hRUP36; hRUP37.
XX OS Homo sapiens.
XX PN WO200242461-A2.
XX PD 30-MAY-2002.
XX PF 26-NOV-2001; 2001WO-US044386.
XX PR 27-NOV-2000; 2000US-0253404P.
XX PR 12-DEC-2000; 2000US-0255366P.
XX PR 20-FEB-2001; 2001US-0270266P.
XX PR 20-FEB-2001; 2001US-0270286P.
XX PR 06-APR-2001; 2001US-0282032P.
XX PR 06-APR-2001; 2001US-0282356P.
XX PR 06-APR-2001; 2001US-0282358P.
XX PR 06-APR-2001; 2001US-0282365P.
XX PR 14-MAY-2001; 2001US-0290917P.
XX PR 31-JUL-2001; 2001US-0309208P.
XX PA (AREN-) ARENA PHARM INC.
XX PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX DR MPI: 2002-56565/60.
XX DR N-BSD; ABT04873.
XX PT Novel endogenous and non-endogenous versions of G protein-coupled
receptor useful for identification of candidate compounds as receptor

```

```

PT agonists or antagonists for use as therapeutic agents.
XX PS Claim 29; Page 70-72; 84pp; English.
XX CC The present invention provides the protein and coding sequences of
XX CC several human G-protein coupled receptors (GPCRs). These can be used in
XX CC the identification of candidate compounds as receptor agonists or inverse
XX CC agonists having applicability as therapeutic agents. The present sequence
XX CC is a GPCR protein of the invention
SQ Sequence 353 AA;
Query Match 100.0%; Score 1872; DB 5; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.8e-200;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSLSWSPSGACGFGVPPVYYSLLCLGLPANILTVIILSQLVARRQK 60
DB 1 MEHTAHLAANSLSWSPSGACGFGVPPVYYSLLCLGLPANILTVIILSQLVARRQK 60
QY 61 SSYNYLLAALAAADILVLFVIVDFLEDFILNQMPQVPDKIEVLEFSSIHHSIMWTV 120
DB 61 SSYNYLLAALAAADILVLFVIVDFLEDFILNQMPQVPDKIEVLEFSSIHHSIMWTV 120
QY 121 PLTIDRYAVCHPLKHTVSPARTKVIYSVYTTCPLTSPYWMWPNWTEYISTSVH 180
DB 121 PLTIDRYAVCHPLKHTVSPARTKVIYSVYTTCPLTSPYWMWPNWTEYISTSVH 180
QY 181 HVLWICFTVYLVPCSIFFILNSIIYKLRKSNPFLRGYSTGKTALILFTTSTPATL 240
DB 181 HVLWICFTVYLVPCSIFFILNSIIYKLRKSNPFLRGYSTGKTALILFTTSTPATL 240
QY 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFRMAAATL 300
DB 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFRMAAATL 300
QY 301 KAFKCKQKOPVQFTYNNHFSITSSPWSPANSHCIKMLVYQYDNGKRPKVSP 353
DB 301 KAFKCKQKOPVQFTYNNHFSITSSPWSPANSHCIKMLVYQYDNGKRPKVSP 353

RESULT 3
ABB79517 ID ABB79517 standard; protein; 353 AA.
XX AC ABB79517;
XX DT 23-SEP-2002 (first entry)
XX DE Human chemokine-like receptor.
XX KM Chemokine-like receptor; G-protein coupled receptor; receptor; human;
XX KW HIV infection; cardiovascular disease; asthma;
XX KW chronic obstructive pulmonary disease; ashiama;
XX KW vasotrophic; hypotensive; antiarrhythmic; thrombolytic; anticoagulant;
XX KW antiinflammatory; antiallergic; immunomodulator; gene therapy.
XX OS Homo sapiens.
XX PN WO200248358-A2.
XX PD 20-JUN-2002.
XX PF 12-DEC-2001; 2001WO-EP014571.
XX PR 14-DEC-2000; 2000US-0255150P.
XX PR 02-APR-2001; 2001US-0280110P.
XX PR 21-JUN-2001; 2001US-0299474P.
XX PA (FARB) BAYER AG.
XX PI Smolyar A, Zhu Z, Encinas J, Watanabe S, Okigami H;
XX PT

```

DR WPI: 2002-547858/58.
 DR N-PSDB; ABBN4269, ABBN4273.
 XX New isolated polynucleotide encoding a chemokine-like receptor
 PT polypeptide for treating e.g. asthma, myocardial infarction, human
 PT immunodeficiency virus infection, or chronic obstructive pulmonary
 PT disease.
 XX
 XX Claim 25, Fig 2, 114pp, English.
 XX
 XX The present sequence is the protein sequence of a novel human chemokine-
 CC like receptor of 353 amino acids. The chemokine-like receptor has 7
 CC putative transmembrane domains, consistent with the structure of a G-
 CC protein coupled receptor. Its closest human homologue is C-C chemokine
 CC receptor 3. The novel receptor is expressed at low levels in most
 CC tissues. It is expressed at a high level in phytohemagglutinin-
 CC stimulated CD8+ cells, but in none of the other immune cells tested. It
 CC may act as a receptor of chemottractant molecules on activated
 CC lymphocytes and be involved in cell trafficking and homing to sites of
 CC infection, inflammation or tissue injury. Regulation of activity of the
 CC novel receptor can therefore be used to treat cardiovascular,
 CC immunological and inflammatory diseases, including asthma and chronic
 CC obstructive pulmonary disease (COPD). The receptor may also be a target
 CC for viruses that reside in the nervous system. Regulating the binding of
 CC ligands, e.g. chemottractant molecules or virus particles, to the
 CC receptor can therefore be used to modulate the immune response to inhibit
 CC viral infections, including HIV infection. A claimed method of reducing
 CC activity of the receptor involves contacting a cell with a reagent
 CC (preferably an antibody, antisense oligonucleotide or ribozyme) to a
 CC product (preferably RNA or a polypeptide) encoded by a polynucleotide
 CC encoding the human chemokine-like receptor in vitro or in vitro. A claimed
 CC method of treating a chemokine-like receptor dysfunction related disease
 CC selected from HIV infection, a cardiovascular disorder, asthma or COPD
 CC uses a reagent that modulates a function of the receptor
 CC
 XX
 XX Sequence 353 AA:
 SQ
 Query Match 100.0%; Score 1872; DB 5; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5.8e-200;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX
 DE Human AXOR-57 protein.
 XX
 XX Antibacterial; antifungal; antiviral; antiprotosol; analgesic;
 KW cytostatic; antidiabetic; anorectic; anabolic; antiallergic; antagonist;
 KW antiparkinsonian; cardiac; hypotensive; hypertensive; nephroretropic;
 KW osteopathic; antiangiatic; anticancer; antiallergic; antimigraine; agonist;
 KW antihemetic; nootropic; tranquilizer; neuroprotective; antidepressant;
 KW vaccine; gene therapy; G-protein coupled receptor; receptor; infection;
 KW pain; cancer; diabetes; obesity; anorexia; bulimia; asthma; stroke;
 KW Parkinson's disease; osteoporosis; angina pectoris; ulcer; allergy;
 KW vomiting; schizophrenia; depression; dementia; Huntington's disease.
 XX
 OS Homo sapiens.
 XX
 PN GB365009-A.
 XX
 PD 13-FEB-2002.
 XX
 PF 11-APR-2001, 2001GB-00009018.
 XX
 PR 11-APR-2001, 2001GB-00009018.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Eshourbagy N, Gattu M, Michalovich D, Shabon U;
 DR WPI: 2003-203569/20.
 DR N-PSDB; ACC44115.
 XX
 PS Claim 1; Page 27-28; 32pp; English.
 XX
 XX The invention relates to the isolation of a novel G-protein coupled (7TM)
 CC receptor AXOR 57 (this sequence) or a DNA fragment encoding a polypeptide
 CC having at least 95% identity with AXOR-57 protein. The protein has been
 CC shown to have homology to the human chromosome 16 clone C17B-H1.036A2
 CC (AC008785). The protein and polynucleotides encoding the protein are used
 CC in screening for compounds that stimulate or inhibit the function or
 CC level of the polypeptide or polynucleotides, such as, agonists and
 CC antagonists. The protein, polynucleotides, and antibodies to the protein
 CC are used in diagnostic kits, to diagnose a disease. Polynucleotides
 CC encoding the protein are used for chromosome localization studies, or for
 CC tissue expression studies. The protein and nucleic acids encoding the
 CC protein are used in vaccines for treating diseases such as, bacterial,
 CC fungal, protozoal, and viral infections, pain, cancer, diabetes, obesity,
 CC anorexia, bulimia, asthma, Parkinson's disease, acute heart failure,
 CC hypotension, hypertension, urinary retention, osteoporosis, angina
 CC pectoris, myocardial infection, stroke, ulcer, allergy, benign prostatic
 CC hypertrophy, migraine, vomiting, anxiety, schizophrenia, depression,
 CC delirium, dementia, severe mental retardation, or Huntington's disease
 CC
 XX
 XX Sequence 353 AA:
 SQ
 Query Match 100.0%; Score 1872; DB 6; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5.8e-200;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 PLTIDRYAVCHPLKHTVSPARTRKIVSVYTICFTLSIPYWMWNIWTEDIYSTSVH 180
QY 181 HTLIMHCTTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTATLFTITSIFATL 240
Db 181 HTLIMHCTTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTATLFTITSIFATL 240
QY 241 WAPRIIMILYHLYGAPIQNRMLVHMSDIANNLALNTAINFELYCFISKFRFMAAATL 300
Db 241 WAPRIIMILYHLYGAPIQNRMLVHMSDIANNLALNTAINFELYCFISKFRFMAAATL 300
QY 301 KAFKCKQKQPVQFYTNHNSITSSPWISPA NSHCKMLVYQYDKGKRPVKSP 353
Db 301 KAFKCKQKQPVQFYTNHNSITSSPWISPA NSHCKMLVYQYDKGKRPVKSP 353

RESULT 5

ABR44437
ID ABR44437 standard; protein; 353 AA.

AC ABR44437;

DT 25-JUL-2003 (first entry)

DE Human G protein-coupled receptor #SEQ ID 2.

KM Human; G protein-coupled receptor; anorectic; eating disorder; obesity.

OS Homo sapiens.

PN WO2003027142-A1.

PD 03-APR-2003.

PF 19-SEP-2002; 2002WO-JP009626.

PR 21-SEP-2001; 2001JP-00288278.

PA (YAMA) YAMANOUCHI PHARM CO LTD.

PI Matsumoto S, Takasaki J, Kurama T, Satto T, Kamohara M, Soga T,
PI Hiyama H;

DR WPI; 2003-333291/31.

DR N-PSDB; ACC71785.

PT New G protein-coupled receptor and encoded gene, useful in screening
PT preventives or remedies for eating disorders or obesity.

PS Claim 1; Page 34-35; 48pp; Japanese.

CC The invention relates to a novel G protein-coupled receptor. The protein
CC and its encoded gene are useful for screening preventives or remedies for
CC eating disorders or obesity. The current sequence represents a G protein
CC coupled receptor sequence

XX Sequence 353 AA;

Query Match Best Local Similarity 100.0%; Score 1872; DB 6; Length 353;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSLSWMSPGSACGIGFVPVYYYSLLCLGLPANTLTVIILSQVAROK 60
Db 1 MEHTAHLAANSLSWMSPGSACGIGFVPVYYYSLLCLGLPANTLTVIILSQVAROK 60
QY 61 SSYNVLLMAADIVLFFIVFVDFLLDFILNQMPQVDPKIIIEVLFSSIHISIMTV 120
Db 61 SSYNVLLMAADIVLFFIVFVDFLLDFILNQMPQVDPKIIIEVLFSSIHISIMTV 120
QY 121 PLTIDRYAVCHPLKHTVSPARTRKIVSVYTICFTLSIPYWMWNIWTEDIYSTSVH 180
Db 121 PLTIDRYAVCHPLKHTVSPARTRKIVSVYTICFTLSIPYWMWNIWTEDIYSTSVH 180

QY 181 HTLIMHCTTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTATLFTITSIFATL 240
Db 181 HTLIMHCTTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTATLFTITSIFATL 240
QY 241 WAPRIIMILYHLYGAPIQNRMLVHMSDIANNLALNTAINFELYCFISKFRFMAAATL 300
Db 241 WAPRIIMILYHLYGAPIQNRMLVHMSDIANNLALNTAINFELYCFISKFRFMAAATL 300
QY 301 KAFKCKQKQPVQFYTNHNSITSSPWISPA NSHCKMLVYQYDKGKRPVKSP 353
Db 301 KAFKCKQKQPVQFYTNHNSITSSPWISPA NSHCKMLVYQYDKGKRPVKSP 353

RESULT 6

ABB82499
ID ABB82499 standard; protein; 353 AA.

AC ABB82499;

DT 22-JAN-2003 (first entry)

DE Human TGR20 polypeptide.

KM G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory;
KM antidiabetic; nootropic; neuroprotective; antiischemic; antidiabetic; human;
KM antiparkinsonian; antidiabetic; TGR20; receptor.

OS Homo sapiens.

PN WO200277001-A2.

PD 03-OCT-2002.

PF 08-MAR-2002; 2002WO-US007171.

PR 09-MAR-2001; 2001US-00802803.

PR 16-MAR-2001; 2001US-0276649P.

PA (TULIA-) TULARIK INC.

PI Tian H, Zhao J, Chen J, Cutler G;

DR WPI; 2003-018881/01.

DR N-PSDB; ABV73364.

PT New G-protein coupled receptor polypeptides and polynucleotides useful
PT for identifying compounds for treating a TGR-associated disorder, e.g.
PT psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's
PT disease, anemia.

PS Claim 15; Page 63; 87pp; English.

CC The invention relates to G-protein coupled receptor (GPCR) polypeptides
CC and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183,
CC TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are
CC useful for identifying compounds for treating a TGR-associated disorder,
CC such as psoriasis, inflammatory bowel disease, hyperlipidemia,
CC Parkinson's disease, Huntington's disease, anemia, immune and blood
CC disorder, ulcerative colitis, Crohn's disease or spleen enlargement.
CC They are also useful for identifying cells such as kidney, liver,
CC hypothalamus, colon, adipose, or spleen cells, for forensics and
CC paternity determination, diagnosing diseases and examining signal
CC transduction. The present sequence represents a human TGR20 polypeptide

XX Sequence 353 AA;

Query Match Best Local Similarity 100.0%; Score 1872; DB 6; Length 353;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSLSWMSPGSACGIGFVPVYYYSLLCLGLPANTLTVIILSQVAROK 60
Db 1 MEHTAHLAANSLSWMSPGSACGIGFVPVYYYSLLCLGLPANTLTVIILSQVAROK 60

QY 61 SSYNLLAALAAADILVLFPIVDFLEDFILNMOMPOVPDKIEVLEFSSHTSIWTV 120
 DB 61 SSYNLLAALAAADILVLFPIVDFLEDFILNMOMPOVPDKIEVLEFSSHTSIWTV 120
 QY 121 PLTIDRYIAVCHPLKHTVSYPARTRKIVSYITCFELSIPIYMMNPINTEDYISTSVH 180
 DB 121 PLTIDRYIAVCHPLKHTVSYPARTRKIVSYITCFELSIPIYMMNPINTEDYISTSVH 180
 QY 181 HVLWICHFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
 DB 181 HVLWICHFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
 QY 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRPRMTAAATL 300
 DB 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRPRMTAAATL 300
 QY 301 KAFKCKQKQVQVQYPTNHNFSTSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353
 DB 301 KAFKCKQKQVQVQYPTNHNFSTSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353

RESULT 7

ID ADF70485 standard; protein; 591 AA.

AC ADF70485;

DT 12-FEB-2004 (first entry)

DE Orphan receptor ligand-related human protein Segid108.

KM ligand; orphan receptor protein; fusion protein; fluorescent protein;
 KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
 KM GFPuv; Enhanced GFP; EGFP; human.

OS Homo sapiens.

PN MO2003071272-A1.

PD 28-AUG-2003.

PF 21-FEB-2003; 2003MO-JP001901.

PR 22-FEB-2002; 2002JP-00045728.

PR 23-JUL-2002; 2002JP-00213949.

PR 11-OCT-2002; 2002JP-00298237.

XX (TAKE) TAKEDA CHEM IND LTD.

PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;

DR N-PSDB; ADF70587.

PT Transformation of cells with a fusion protein of an orphan receptor
 PT protein with a fluorescent protein useful for identification of ligands
 PT to the orphan receptor.

PS Disclosure; SEQ ID NO 108; 594pp; Japanese.

CC This invention relates to a novel method of identifying ligands to an
 CC orphan receptor protein which comprises transforming cells with DNA
 CC encoding a fusion protein of the orphan receptor with a fluorescent
 CC protein, so that the fusion protein is expressed in the cells (or cell
 CC membranes isolated from them) and contacting the cells with the potential
 CC ligand to be tested. A suitable fluorescent protein for incorporation in
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
 CC identification of ligands binding to an orphan receptor protein.

XX Sequence 591 AA;

Query Match 100.0%; Score 1872; DB 7; Length 591;
 Best Local Similarity 100.0%; Pred. No. 1,2e-199;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSWSGSAAGLGFVYVYSLLCGLPANILTVIILSQLVAROK 60
 DB 1 MEHTHAHLAANSLSWSGSAAGLGFVYVYSLLCGLPANILTVIILSQLVAROK 60
 QY 61 SSYNLLAALAAADILVLFPIVDFLEDFILNMOMPOVPDKIEVLEFSSHTSIWTV 120
 DB 61 SSYNLLAALAAADILVLFPIVDFLEDFILNMOMPOVPDKIEVLEFSSHTSIWTV 120
 QY 121 PLTIDRYIAVCHPLKHTVSYPARTRKIVSYITCFELSIPIYMMNPINTEDYISTSVH 180
 DB 121 PLTIDRYIAVCHPLKHTVSYPARTRKIVSYITCFELSIPIYMMNPINTEDYISTSVH 180
 QY 181 HVLWICHFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
 DB 181 HVLWICHFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
 QY 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRPRMTAAATL 300
 DB 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRPRMTAAATL 300
 QY 301 KAFKCKQKQVQVQYPTNHNFSTSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353
 DB 301 KAFKCKQKQVQVQYPTNHNFSTSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353

RESULT 8

ID AAU10068 standard; protein; 353 AA.

AC AAU10068;

DT 14-FEB-2002 (first entry)

DE Chemokine receptor family related protein, NOV11.

KM NOV; cytostatic; psoriasis; noctropic; neuroprotectant;
 KW cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;
 KW haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;
 KW differentiation; proliferation; haematopoiesis; wound healing;
 KW angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;
 KW haemophilia; allergy; Perked syndrome; skeletal dysplasia;
 KW ischemic injury; neuroepithelial disorder; hepatitis; heart failure;
 KW chemokine receptor; chromosome 1.

OS Homo sapiens.

PI Key Location/Qualifiers

FT Peptide 1..47 /label= Signal_peptide

FT Protein 48..353 /note= "Mature chemokine receptor related protein, NOV11"

PN MO200170978-A2.

XX 27-SEP-2001.

PF 20-MAR-2001; 2001MO-US009093.

PR 20-MAR-2000; 2000US-0190768P.

PR 20-MAR-2000; 2000US-0190835P.

PR 22-MAR-2000; 2000US-0190972P.

PR 24-MAR-2000; 2000US-0191199P.

PR 26-MAR-2000; 2000US-0191477P.

PR 28-MAR-2000; 2000US-0192657P.

PR 28-MAR-2000; 2000US-0192657P.

PR 29-MAR-2000; 2000US-0192836P.

PR 31-MAR-2000; 2000US-0193843P.

XX (CURA-) CURAGEN CORP.
 XX
 PI Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;
 PI Vernet CAM;
 XX
 DR WPI: 2001-639127/73.
 DR N-PSDB; AAS15731.
 XX
 PT Polypeptides and nucleic acids related to chloride channel, insulin-like
 PT growth factor family of proteins, useful for diagnosing and treating
 PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
 XX
 PS Claim 1, Page 44; 151pp; English.

XX The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX
 CC polypeptides are useful for treating pathology associated with NOVX
 CC polypeptide, determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX, identifying agents binding to
 CC NOVX and treatment of disorders associated with altered expression of
 CC members of chloride channel-associated proteins e.g. cystic fibrosis and
 CC congenital myotonia. NOVX proteins are useful in treatment of disorders
 CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,
 CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
 CC in one or more organs (e.g. haemophilia, anaemia), pendred syndrome,
 CC skeletal dysplasia, disorders characterised by altered cell shape,
 CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial
 CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
 CC treatment of disorders of vascular smooth muscle cell differentiation,
 CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are
 CC useful to screen for molecules which inhibit or enhance NOVX activity or
 CC function and are useful as targets for the identifying small molecules,
 CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
 CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV
 CC sequences are also useful for: identifying a cell or tissue type in a
 CC biological sample; amplifying DNA sequences from very small biological
 CC samples e.g. hair or skin or body fluids and as primers and probes to
 CC identify and/or clone NOVX homologues. NOVX proteins are useful
 CC immunogens to generate antibodies to monitor protein levels and modulate
 CC NOVX activity. Cells comprising the nucleic acids are useful for
 CC producing transgenic animals, for studying the function and/or activity
 CC of NOVX protein and identifying and/or evaluating modulators of NOVX
 CC protein activity. This sequence is the NOVX1 amino acid sequence (gene
 CC located on chromosome 1) related to the chemokine receptor family of
 CC proteins, one of 12 NOV polypeptides described in the method of the
 CC invention
 XX
 XX
 SQ Sequence 353 AA;

Query Match 99.8%; Score 1868; DB 4; Length 353;
 Best Local Similarity 99.7%; Pred. No. 1.6e-199;
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTAHNLANSLSWSWSPGACGFGVYVYVYLLCLGIPANLTYIIISQVAPAROK 60
 DB 1 MEHTAHNLANSLSWSWSPGACGFGVYVYVYLLCLGIPANLTYIIISQVAPAROK 60
 QY 61 SSYNVLLAAADIIIVLFYIVFVDFLDFILNMQMPQVPKIIIEVLFSSIHHSIWTV 120
 DB 61 SSYNVLLAAADIIIVLFYIVFVDFLDFILNMQMPQVPKIIIEVLFSSIHHSIWTV 120
 QY 121 PLTIDRYAVCHPLKHTVSPARTKRVISYVITCEFTLSIPLYWMPNITWEDYISVH 180
 DB 121 PLTIDRYAVCHPLKHTVSPARTKRVISYVITCEFTLSIPLYWMPNITWEDYISVH 180
 QY 121 PLTIDRYAVCHPLKHTVSPARTKRVISYVITCEFTLSIPLYWMPNITWEDYISVH 180
 DB 121 PLTIDRYAVCHPLKHTVSPARTKRVISYVITCEFTLSIPLYWMPNITWEDYISVH 180
 QY 181 HVLIIHCFYVYLVCSIFFIINSIIVYKLRKSNFRLRGYSTGKTAAIIFTISIFATL 240
 DB 181 HVLIIHCFYVYLVCSIFFIINSIIVYKLRKSNFRLRGYSTGKTAAIIFTISIFATL 240
 QY 241 WAPRIIMLYHYGAPIONRWLVHMSDIANMLALNTAINFHYCISKRFRMAATL 300
 DB 241 WAPRIIMLYHYGAPIONRWLVHMSDIANMLALNTAINFHYCISKRFRMAATL 300

QY 301 KAFFKOKOPVOPTNHNFSITSSPWISPANSHCTKMLVYQDKXGKPIKSP 353
 DB 301 KAFFKOKOPVOPTNHNFSITSSPWISPANSHCTKMLVYQDKXGKPIKSP 353

RESULT 9
 ADJ87769
 ID ADJ87769 standard; protein; 353 AA.
 XX
 AC ADJ87769;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE G-coupled protein receptor-related protein #57.
 XX

XX novel protein; G-coupled protein receptor-related protein;
 KW cardiomyopathy; atherosclerosis; cell signal processing-related disorder;
 KW metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
 KW Huntington's disease; epilepsy; anxiety; pain; hypercholesterolaemia;
 KW obesity; hypertension; Crohn's disease; systemic lupus erythematosus;
 KW viral infections; bacterial infection; parasitic infection;
 KW hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;
 KW Alzheimer's disease; tuberculous sclerosis; hypercalcaemia; cerebral palsy.
 XX
 OS Unidentified.
 XX
 PN W02002102321-A2.
 XX
 PD 27-DEC-2002.
 XX
 XX 18-JUN-2002; 2002WO-US019522.
 XX
 PF 18-JUN-2001; 2001US-0298994P.
 XX
 PR 18-JUN-2001; 2001US-0299134P.
 PR 04-OCT-2001; 2001US-00972446.
 PR 06-JUN-2002; 2002US-00299134.
 PR 07-JUN-2002; 2002US-00298994.
 XX

PA (CURA-) CURAGEN CORP.
 XX
 XX Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;
 PI Pena CEA, Spaderna SK, Zhong M;
 XX
 DR WPI: 2003-167441/16.
 DR N-PSDB; ADJ87768.
 XX
 XX New MOX polypeptides and polynucleotides, useful in gene therapy,
 FT particularly for treating or preventing e.g. cardiomyopathy,
 FT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
 FT cancer, stroke or pain.
 XX
 PS Claim 1, SEQ ID NO 204; 378pp; English.

XX The invention comprises the amino acid and coding sequences of novel G-
 CC coupled protein receptor-related (MOX) proteins. The DNA and protein
 CC sequences of the invention are useful for treating or preventing a MOX-
 CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
 CC or diabetes. The DNA and protein sequences are also useful for the
 CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
 CC pain, hypercholesterolaemia, obesity, hypertension, Crohn's disease,
 CC systemic lupus erythematosus, viral infections, bacterial infections,
 CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
 CC syndrome, Alzheimer's disease, tuberculous sclerosis, hypercalcaemia, or
 CC cerebral palsy. The present amino acid sequence represents a MOX protein
 CC of the invention.
 XX

Query Match 99.8%; Score 1868; DB 7; Length 353;
 Best Local Similarity 99.7%; Pred. No. 1.6e-199;
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTAHLANSSLSWMSGACGAGFVPPVYYSLLCGLGPANILTVIIISQVBAROK 60
 DB 1 MEHTAHLANSSLSWMSGACGAGFVPPVYYSLLCGLGPANILTVIIISQVBAROK 60
 QY 61 SSYNVLLAALAAADIVLFFIVFDPLLEDFFILNMQMPQVPDKIIEVLEFSSIHITSIMITY 120
 DB 61 SSYNVLLAALAAADIVLFFIVFDPLLEDFFILNMQMPQVPDKIIEVLEFSSIHITSIMITY 120
 QY 121 PLTIDRYIAVCHPLKXHTVSYPARTRKVIYSYITCFLTSIPYMPNIMTEDYISTSYH 180
 DB 121 PLTIDRYIAVCHPLKXHTVSYPARTRKVIYSYITCFLTSIPYMPNIMTEDYISTSYH 180
 QY 181 HVLWTHCFYVLPVPSIFILNSIIVYKLRKSNRLNGYSTGKTTALFTTISIFATL 240
 DB 181 HVLWTHCFYVLPVPSIFILNSIIVYKLRKSNRLNGYSTGKTTALFTTISIFATL 240
 QY 241 MAPRIIMILYHAGAPIQRMLVHIMSDIANMLALNTAINPFLYCFISKRPTMAATL 300
 DB 241 MAPRIIMILYHAGAPIQRMLVHIMSDIANMLALNTAINPFLYCFISKRPTMAATL 300
 QY 301 KAFKQKQKQPVQFYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIKVP 353
 DB 301 KAFKQKQKQPVQFYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIKVP 353
 RESULT 10
 ADI79325
 ID ADI79325 standard; protein; 353 AA.
 XX
 AC ADI79325;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE NOV11 protein sequence, SEQ ID 22.
 XX
 KM Cytostatic; Immunosuppressive; Anti-allergic; Antimicrobial; Vasotropic;
 KM Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;
 KM Ophthalmological; Antipsoriatic; Neuroprotective; Nocrotic;
 KM Antiarteriosclerotic; Hypotensive; Cardiant; Cerebroprotective;
 KM Gene Therapy; NOVX; human; cancer; myelogenous leukemia;
 KM congenital neonatal autoimmune thrombocytopaenia; immunological disorder;
 KM allergy; infection; asthma; lung disease; reproductive disorder;
 KM haemangioma; deafness; liver cirrhosis; hepatitis C; gastric disorder;
 KM diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;
 KM hypertension; stroke; heart failure; chromosome 1; NOV11;
 KM chemokine receptor.
 XX
 OS Homo sapiens.
 XX
 PN MO2004009635-A2.
 XX
 PD 29-JAN-2004.
 XX
 PF 04-OCT-2001; 2001MO-US031292.
 XX
 PR 20-MAR-2001; 2001US-00813432.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Taupler RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;
 PI Vernet CAM;
 XX
 DR MPI; 2004-123380/12.
 XX
 PS N-PSDB; ADI79324.
 XX
 PS Claim 1; Page 44; 15pp; English.
 XX
 CC The present invention relates to novel NOVX proteins and their coding
 CC sequences (ADI79304-ADI79327). The sequences are useful for the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease associated with the protein, or for diagnosing and treating
 CC disorders associated with the NOVX protein, such as cancer, myelogenous
 CC leukemia, congenital neonatal autoimmune thrombocytopaenia,

CC immunological disorders, allergy and infection, asthma, lung diseases,
 CC reproductive disorders, male and female reproductive diseases,
 CC haemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,
 CC diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,
 CC hypertension, stroke and heart failure. NOV11 represents a new member of
 CC the chemokine receptor family and the gene is located on chromosome 1.
 XX
 SQ Sequence 353 AA:
 Query Match 99.8%; Score 1868; DB 8; Length 353;
 Best Local Similarity 99.7%; Pred. No. 1.6e-199;
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MEHTAHLANSSLSWMSGACGAGFVPPVYYSLLCGLGPANILTVIIISQVBAROK 60
 DB 1 MEHTAHLANSSLSWMSGACGAGFVPPVYYSLLCGLGPANILTVIIISQVBAROK 60
 QY 61 SSYNVLLAALAAADIVLFFIVFDPLLEDFFILNMQMPQVPDKIIEVLEFSSIHITSIMITY 120
 DB 61 SSYNVLLAALAAADIVLFFIVFDPLLEDFFILNMQMPQVPDKIIEVLEFSSIHITSIMITY 120
 QY 121 PLTIDRYIAVCHPLKXHTVSYPARTRKVIYSYITCFLTSIPYMPNIMTEDYISTSYH 180
 DB 121 PLTIDRYIAVCHPLKXHTVSYPARTRKVIYSYITCFLTSIPYMPNIMTEDYISTSYH 180
 QY 181 HVLWTHCFYVLPVPSIFILNSIIVYKLRKSNRLNGYSTGKTTALFTTISIFATL 240
 DB 181 HVLWTHCFYVLPVPSIFILNSIIVYKLRKSNRLNGYSTGKTTALFTTISIFATL 240
 QY 241 MAPRIIMILYHAGAPIQRMLVHIMSDIANMLALNTAINPFLYCFISKRPTMAATL 300
 DB 241 MAPRIIMILYHAGAPIQRMLVHIMSDIANMLALNTAINPFLYCFISKRPTMAATL 300
 QY 301 KAFKQKQKQPVQFYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIKVP 353
 DB 301 KAFKQKQKQPVQFYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIKVP 353
 RESULT 11
 AD056004
 ID AD056004 standard; protein; 353 AA.
 XX
 AC AD056004;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Human NOV11 polypeptide.
 XX
 KM human; cancer; obesity; diabetic nephropathy; acute pancreatitis; stroke;
 KM multiple sclerosis.
 XX
 OS Homo sapiens.
 XX
 PN US2004058862-A1.
 XX
 PD 25-MAR-2004.
 XX
 PF 18-SEP-2002; 2002US-00246583.
 XX
 PR 18-SEP-2002; 2002US-00246583.
 XX
 PA (MAU/) MAJUMDER K.
 XX
 PI Majumder K;
 XX
 DR MPI; 2004-268835/25.
 XX
 PS N-PSDB; AD056003.
 XX
 PT Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic
 PT nephropathy, acute pancreatitis, strokes and multiple sclerosis.
 XX
 PS Disclosure; Page 28; 87pp; English.
 XX

CC The invention relates to novel isolated NOVX nucleic acids and encoded
 CC polypeptides. The nucleic acids, polypeptides and antibodies raised
 CC against the polypeptides are useful for preventing or treating diseases
 CC associated with aberrant NOVX expression or activity e.g., cancer,
 CC obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple
 CC sclerosis. The present sequence represents a NOVX polypeptide of the
 CC invention.

CC Sequence 353 AA;

Query Match 99.8%; Score 1868; DB 8; Length 353;
 Best Local Similarity 99.7%; Pred. No. 1.6e-199;
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSLSWSPGSAAGLGFVPPVYYSLLCLGIPANILTVIISQVARRQK 60
 DB 1 MEHTAHLAANSLSWSPGSAAGLGFVPPVYYSLLCLGIPANILTVIISQVARRQK 60
 QY 61 SSYNVLLAALAAADILVLFVIVDFLLEDFILNMQMPQVDPKIIIVLEFSSHTSIWTV 120
 DB 61 SSYNVLLAALAAADILVLFVIVDFLLEDFILNMQMPQVDPKIIIVLEFSSHTSIWTV 120
 QY 121 PLTIDRYAVCHPLKHTVSYPARTRKYIVSYITTCFLTSPYYWPNWITEDYISTSVH 180
 DB 121 PLTIDRYAVCHPLKHTVSYPARTRKYIVSYITTCFLTSPYYWPNWITEDYISTSVH 180
 QY 181 HVLIMHCFVTVLPGCSIFFLINSIIYKLRKSNFRLRGYSTGKTALFTTSTFATL 240
 DB 181 HVLIMHCFVTVLPGCSIFFLINSIIYKLRKSNFRLRGYSTGKTALFTTSTFATL 240
 QY 241 WAPRIIMILYHLYGAPIONRWLVHMSDIANMLALNTAINPFLYCFISKRFTMAATL 300
 DB 241 WAPRIIMILYHLYGAPIONRWLVHMSDIANMLALNTAINPFLYCFISKRFTMAATL 300
 QY 301 KAEFFCKQKQPVQFYTHNPSITSSPMTSPANSHCITMLVQYKNGKPIKVP 353
 DB 301 KAEFFCKQKQPVQFYTHNPSITSSPMTSPANSHCITMLVQYKNGKPIKVP 353

RESULT 12
 ADO28954
 ID ADO28954 standard; protein; 353 AA.

AC ADO28954;

DT 29-JUL-2004 (first entry)

DE Human novel GPCR PGR3, SEQ ID NO:53.

XX G Protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; antitumor;
 KW cytotoxic; antineoplastic; vasotropic; antiangiogenic; antiarhythmic;
 KW CNS; central nervous system; respiratory; antiatherosclerotic;
 KW virologic; hepatotropic; antibacterial; antidiabetic; antidiabetic;
 KW dermatological; antitumor; antihypertensive; antiepileptic;
 KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KW receptor.

XX Homo sapiens.

XX WO2004040000-A2.

XX 13-MAY-2004.

XX 09-SEP-2003; 2003WO-US028226.

PR 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX (PRIM-) PRIMAL INC.
 XX Galtanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F,
 PI Madison L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H,
 DR WPI; 2004-390329/36.
 XX N-PSDB; ADO28955.
 PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PS pectoris, Parkinson's disease.

PS Claim 1, SEQ ID NO 53; 542PP; English.

XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anemia or leukemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancer). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 353 AA;

Query Match 99.8%; Score 1868; DB 8; Length 353;
 Best Local Similarity 99.7%; Pred. No. 1.6e-199;
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSLSWSPGSAAGLGFVPPVYYSLLCLGIPANILTVIISQVARRQK 60
 DB 1 MEHTAHLAANSLSWSPGSAAGLGFVPPVYYSLLCLGIPANILTVIISQVARRQK 60
 QY 61 SSYNVLLAALAAADILVLFVIVDFLLEDFILNMQMPQVDPKIIIVLEFSSHTSIWTV 120
 DB 61 SSYNVLLAALAAADILVLFVIVDFLLEDFILNMQMPQVDPKIIIVLEFSSHTSIWTV 120
 QY 121 PLTIDRYAVCHPLKHTVSYPARTRKYIVSYITTCFLTSPYYWPNWITEDYISTSVH 180
 DB 121 PLTIDRYAVCHPLKHTVSYPARTRKYIVSYITTCFLTSPYYWPNWITEDYISTSVH 180
 QY 181 HVLIMHCFVTVLPGCSIFFLINSIIYKLRKSNFRLRGYSTGKTALFTTSTFATL 240
 DB 181 HVLIMHCFVTVLPGCSIFFLINSIIYKLRKSNFRLRGYSTGKTALFTTSTFATL 240
 QY 241 WAPRIIMILYHLYGAPIONRWLVHMSDIANMLALNTAINPFLYCFISKRFTMAATL 300

DB 241 MARRIIMILYHVGAPIQRMVHIMSDIANMLALNTAINFLYCFISKRRTMAAATL 300
 QY 301 KAFKCKQKQPVQFYTNHNFSTSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353
 DB 301 KAFKCKQKQPVQFYTNHNFSTSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353

RESULT 13
 AAU10067
 ID AAU10067 standard; protein; 372 AA.
 AC AAU10067;
 DT 14-FEB-2002 (first entry)
 DE Chemokine receptor family related protein, NOV10.
 XX
 XX NOV; cytostatic; psoriasis; noctropia; neuroprotectant;
 KM cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;
 KM hemostatic;therosclerosis; gene therapy; neurogenesis; motility;
 KM differentiation; proliferation; haematopoiesis; wound healing;
 KM angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;
 KM hemophilis; allergy; Pendred syndrome; skeletal dysplasia;
 KM ischemic injury; neuroepithelial disorder; hepatitis; heart failure;
 KM chemokine receptor; chromosome 1.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..47
 FT /label= Signal_peptide
 FT Protein 48..372
 FT /note= "Mature chemokine receptor related protein, NOV10"
 XX
 PN WO200170978-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 20-MAR-2001; 2001WO-US09093.
 XX
 PR 20-MAR-2000; 2000US-0190768P.
 PR 20-MAR-2000; 2000US-0190835P.
 PR 22-MAR-2000; 2000US-0190972P.
 PR 22-MAR-2000; 2000US-0191199P.
 PR 24-MAR-2000; 2000US-0191947P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192664P.
 PR 28-MAR-2000; 2000US-0192665P.
 PR 28-MAR-2000; 2000US-0192884P.
 PR 29-MAR-2000; 2000US-0192836P.
 PR 31-MAR-2000; 2000US-0193843P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Taupier RJ, Meijumder K, Spederna SK, Smithson G, Mezes PS;
 PI Vernet CAM;
 DR MPI; 2001-639127/73.
 DR N-PSDB; AAS15730.
 XX
 PT Polypeptides and nucleic acids related to chloride channel, insulin-like
 PT growth factor family of proteins, useful for diagnosing and treating
 PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
 XX
 PS Claim 1; Page 42; 151pp; English.
 XX
 CC The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX
 CC polypeptides are useful for treating pathology associated with NOVX
 CC polypeptide, determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX, identifying agents binding to
 CC NOVX and treatment of disorders associated with altered expression of
 CC members of chloride channel-associated proteins e.g. cystic fibrosis and

CC congenital myotonia. NOVX proteins are useful in treatment of disorders
 CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,
 CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
 CC in one or more organs (e.g. haemophilia, anaemia), Pendred syndrome,
 CC skeletal dysplasias, disorders characterised by altered cell shape,
 CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial
 CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
 CC treatment of disorders of vascular smooth muscle cell differentiation,
 CC (e.g. heart failure, strokes). NOVX nucleic acids and polypeptides are
 CC useful to screen for molecules which inhibit or enhance NOVX activity or
 CC function and are useful as targets for the identifying small molecules,
 CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
 CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV
 CC sequences are also useful for: identifying a cell or tissue type in a
 CC biological sample; amplifying DNA sequences from very small biological
 CC samples e.g. hair or skin or body fluids and as primers and probes to
 CC identify and/or clone NOVX homologues. NOVX proteins are useful
 CC immunogens to generate antibodies to monitor protein levels and modulate
 CC NOVX activity. Cells comprising the nucleic acids are useful for
 CC producing transgenic animals, for studying the function and/or activity
 CC of NOVX protein and identifying and/or evaluating modulators of NOVX
 CC protein activity. This sequence is the NOV10 amino acid sequence (gene
 CC located on chromosome 1) related to the chemokine receptor family of
 CC proteins, one of 12 NOV polypeptides described in the method of the
 CC invention
 XX
 SO Sequence 372 AA;
 XX
 Query Match 99.2%; Score 1857; DB 4; Length 372;
 Best Local Similarity 100.0%; Pred. No. 3e-198;
 Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEHTAHMLAANSLSLWSGSGACGFGVYVYYSLLCLGAPANTITVYIISQVARROK 60
 DB 1 MEHTAHMLAANSLSLWSGSGACGFGVYVYYSLLCLGAPANTITVYIISQVARROK 60
 QY 61 SSYNYTLAAADIVLFPFVDFLEBDFILNMQPVDPKIEVLEBSIHTSIWTV 120
 DB 61 SSYNYTLAAADIVLFPFVDFLEBDFILNMQPVDPKIEVLEBSIHTSIWTV 120
 QY 121 PLTIDRYIIVCHPLKHTVSPARTKVIIVYITCFELTSDPYMMPNITWEDYISTSVH 180
 DB 121 PLTIDRYIIVCHPLKHTVSPARTKVIIVYITCFELTSDPYMMPNITWEDYISTSVH 180
 QY 181 HVLWTHCTVYLVPCSIPIFINSTIIVYLKRSNFRLEGGYSTGTTAIFITTSIPATL 240
 DB 181 HVLWTHCTVYLVPCSIPIFINSTIIVYLKRSNFRLEGGYSTGTTAIFITTSIPATL 240
 QY 241 WAPRIIMILYHVGAPIQRMVHIMSDIANMLALNTAINFLYCFISKRRTMAAATL 300
 DB 241 WAPRIIMILYHVGAPIQRMVHIMSDIANMLALNTAINFLYCFISKRRTMAAATL 300
 QY 301 KAFKCKQKQPVQFYTNHNFSTSSPMISPANSHCIKMLVYQYDKNGKPIK 350
 DB 301 KAFKCKQKQPVQFYTNHNFSTSSPMISPANSHCIKMLVYQYDKNGKPIK 350

RESULT 14
 AAB18645
 ID AAB18645 standard; protein; 372 AA.
 AC AAB18645;
 DT 17-MAY-2002 (first entry)
 DE Human G-protein coupled receptor (GPRC-6).
 XX
 XX Human; G-protein coupled receptor, GPRC-6; cell proliferative disorder;
 KM neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;
 KM metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;
 KM Pick's disease; Huntington's disease; Parkinson's disease; hypertension;
 KM atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;
 KM osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;

CC	GCRC	expression	The present sequence is human GCRC-6
XX	Sequence	372 AA	
XX	Query Match	99.2%; Score 1857; DB 5; Length 372;	
	Best Local Similarity	100.0%; Pred. No. 3e-199;	
	Matches	350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	MEHTAHHLAANSSLSWSPGSACGLGFPVYVYSLLCGLPANILTVIILSQLVARROK	60
Db	1	MEHTAHHLAANSSLSWSPGSACGLGFPVYVYSLLCGLPANILTVIILSQLVARROK	60
QY	61	SSNYNLALAAADILVEPIFVDFLEDEFLMMQMPQVDPDKILEVLEPSSHTSITWTV	120
Db	61	SSNYNLALAAADILVEPIFVDFLEDEFLMMQMPQVDPDKILEVLEPSSHTSITWTV	120
QY	121	PLTIRRYAVCHPLKHTVTSYPARTKRVISVYITGFLSIPYMMPNITWEDYSTSVH	180
Db	121	PLTIDRYAVCHPLKHTVTSYPARTKRVISVYITGFLSIPYMMPNITWEDYSTSVH	180
QY	181	HVLIMHCFYVYLPCISIFFLINSIIYVKKRKSNNRLGSGTKTALLFTTSTIFATL	240
Db	181	HVLIMHCFYVYLPCISIFFLINSIIYVKKRKSNNRLGSGTKTALLFTTSTIFATL	240
QY	241	WAPRIIMTLVHLVAPIQNFMVLVHNSDIANNMALLNTAINPFLYCPISKRFRMAAATL	300
Db	241	WAPRIIMTLVHLVAPIQNFMVLVHNSDIANNMALLNTAINPFLYCPISKRFRMAAATL	300
QY	301	KAFKCKQKQVQFYTNHNFISITSSPISPANSHCIKMLVYQVDKNGKPIK	350
Db	301	KAFKCKQKQVQFYTNHNFISITSSPISPANSHCIKMLVYQVDKNGKPIK	350
RESULT 15			
ABR62521			
ID	ABR62521	standard; protein; 372 AA.	
XX	ABR62521;		
AC	06-NOV-2003	(first entry)	
XX			
DE	Human G-protein coupled receptor HGRBMY34.		
XX			
XX	HGRBMY34; G-protein coupled receptor; receptor; GPCR-p14; GPCR-145;		
KM	human; neuroprotective; nootropic; transnillizer; antimigraine;		
KM	neuroleptic; antiemetic; antidepressant; anticonvulsant; antiparkinsonian;		
KM	cytostatic; cardiant; hypotensive; antianginal; analgesic; anorectic;		
KM	anti-HIV; antisthmatic; osteopathic; utropanic; antulcer; antiallergic;		
XX	gene therapy.		
OS	Homo sapiens.		
XX			
FX			
Key	Location/Qualifiers		
FT	16..36		
FT	/label= TM1		
FT	/note= "transmembrane domain 1"		
FT	65..87		
FT	/label= TM2		
FT	/note= "transmembrane domain 2"		
FT	109..131		
FT	/label= TM3		
FT	/note= "transmembrane domain 3"		
FT	148..166		
FT	/label= TM4		
FT	/note= "transmembrane domain 4"		
FT	182..208		
FT	/label= TM5		
FT	/note= "transmembrane domain 5"		
FT	227..249		
FT	/label= TM6		
FT	/note= "transmembrane domain 6"		
FT	269..288		
FT	/label= TM7		
Domain			

```

FT      /note="transmembrane domain 7"
XX      MO2003050256-A2.
PN      19-JUN-2003.
XX      06-DEC-2002; 2002MO-US039290.
XX      06-DEC-2001; 2001US-0338371P.
XX      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX      Feder JN, Gopal S, Mincier GA, Ramanathan CS;
XX      WPI, 2003-577295/54.
XX      N-PSDB; ACF05275.
XX      New nucleic acid molecule encoding a human G-protein coupled receptor,
PT      HGPBMY34, useful for diagnosing, preventing or treating diseases
PT      involving the receptor, for example Parkinson's disease, dementia,
PT      asthma, hypertension or cancer.
XX      Claim 5; Fig 1A-B; 112pp; English.
XX      The present sequence is the protein sequence of human HGPBMY34, a newly
CC      identified G-protein coupled receptor (GPCR) belonging to the group of
CC      Class A GPCRs and showing homology to the Pfam model 7 transmembrane
CC      receptor, rhodopsin family. HGPBMY34, also referred to as GPCR-P14
CC      and/or GPCR-145, is highly expressed in brain (especially in the
CC      amygdala, caudate nucleus, corpus callosum, hippocampus, thalamus,
CC      substantia nigra), spinal cord and pituitary, indicating an association
CC      in neurological systems and conditions. It is also expressed in the bone
CC      marrow and testis. The invention provides HGPBMY34 polynucleotides,
CC      polypeptides and antibodies, expression vectors, host cells and antisense
CC      molecules, methods for screening for modulators of HGPBMY34 activity
CC      and/or function, and methods for diagnosing, treating, preventing and
CC      screening for disorders and diseases associated with abnormal HGPBMY34
CC      activity, including: a disorder related to aberrant G-protein coupled
CC      signaling; a disorder related to aberrant cell cycle regulation;
CC      neurological disorders; anxiety; headache; migraine; schizophrenia; manic
CC      depression; delirium; dementia; severe mental retardation and
CC      dyskinesias, such as Huntington's disease or Gilles de la Tourette's
CC      syndrome; Parkinson's disease; brain disorders; spinal cord disorders;
CC      affective disorders; neoplastic disorders; cardiovascular disorders;
CC      acute heart failure; hypotension; hypertension; angina pectoris;
CC      myocardial infarction; an immunological disorder; immune-related
CC      disorders; endocrinal diseases; growth disorders; neuropathic pain;
CC      obesity; anorexia; HIV infections; cancers; bulimia; asthma; osteoporosis
CC      ; psychosis; metabolic disorders; pituitary disorders; urinary retention;
CC      ulcers; allergies; or benign prostatic hypertrophy (all claimed)
XX      SQ
XX      Sequence 372 AA:
XX
XX      Query Match          99.2%; Score 1857; DB 6; Length 372;
XX      Best Local Similarity 100.0%; Pred. No. 3e-198;
XX      Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY      1 MEHTHAHLAANSSLSWSPGACGLGFVPPVYVYSLLLCGLPANILTVIILSQLVAROK 60
DB      1 MEHTHAHLAANSSLSWSPGACGLGFVPPVYVYSLLLCGLPANILTVIILSQLVAROK 60
QY      61 SSVNYLLAALAAADILVLFVFPVDFLEDFILNMQMPQVPDKIIEVLEFSSHTSIWTV 120
DB      61 SSVNYLLAALAAADILVLFVFPVDFLEDFILNMQMPQVPDKIIEVLEFSSHTSIWTV 120
QY      121 PLTIDRYIANCHPLKHYTVSPARTKVIYSVYTCTPLTSPYVWMPNMTEDYISTSVH 180
DB      121 PLTIDRYIANCHPLKHYTVSPARTKVIYSVYTCTPLTSPYVWMPNMTEDYISTSVH 180
QY      181 HVLIIWICFTVYLVPSCIFPILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIPATL 240
DB      181 HVLIIWICFTVYLVPSCIFPILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIPATL 240

```

```

QY      241 WAPRIIMILYHLYGAPIONRWLVHMSDIPANMLALNTAINFPLYCFISKRFRMAATL 300
DB      241 WAPRIIMILYHLYGAPIONRWLVHMSDIPANMLALNTAINFPLYCFISKRFRMAATL 300
QY      301 KAFKCOQKQPVQPYTNHNSITSSPWIISPANSHCINMLVYQYDKNGKPIK 350
DB      301 KAFKCOQKQPVQPYTNHNSITSSPWIISPANSHCINMLVYQYDKNGKPIK 350

```

Search completed: December 14, 2004, 20:23:08
 Job time : 59.6035 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 14, 2004, 20:25:14 ; Search time 16.4665 Seconds

(Without alignments)
1421.692 Million cell updates/sec

Title: US-09-995-225B-16

Perfect score: 1872
Sequence: 1 MHTTAHLAANSSLSWSPG.....CIKMLVYQDKXKPIKVP 353Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 6631800 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/1aa/6C_COMB.pep.*
6: /cgn2_6/prodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	241	12.9	393	1	US-07-629-1041-3 Sequence 3, App1
2	238.5	12.7	398	2	US-08-288-663A-1 Sequence 1, App1
3	232	12.4	415	4	US-09-545-844-2 Sequence 2, App1
4	231.5	12.4	355	1	US-08-012-988A-2 Sequence 2, App1
5	231.5	12.4	355	1	US-08-450-933A-5 Sequence 5, App1
6	231.5	12.4	355	3	US-08-446-669-5 Sequence 5, App1
7	231.5	12.4	355	3	US-09-239-938-1 Sequence 1, App1
8	231.5	12.4	355	4	US-09-886-119A-14 Sequence 14, App1
9	231.5	12.4	355	4	US-10-039-659A-13 Sequence 13, App1
10	231.5	12.4	355	4	US-09-961-068-1 Sequence 1, App1
11	231.5	12.4	355	4	US-09-625-973-5 Sequence 5, App1
12	231.5	12.4	355	4	US-09-960-547-1 Sequence 1, App1
13	231.5	12.4	355	5	PCT-US95-00476-5 Sequence 5, App1
14	226.5	12.1	398	2	US-08-288-663A-15 Sequence 15, App1
15	223.5	11.9	355	4	US-08-833-752-9 Sequence 9, App1
16	223.5	11.9	355	4	US-09-938-719-9 Sequence 9, App1
17	223	11.9	352	4	US-09-029-027B-2 Sequence 2, App1
18	222.5	11.9	347	1	US-08-118-270-47 Sequence 47, App1
19	222.5	11.9	347	1	PCT-US93-08528-47 Sequence 47, App1
20	219	11.7	353	5	US-08-118-270-45 Sequence 45, App1
21	219	11.7	353	5	PCT-US93-08528-45 Sequence 45, App1
22	219	11.7	357	1	US-08-149-893A-4 Sequence 4, App1
23	219	11.7	357	2	US-08-911-245-4 Sequence 4, App1
24	219	11.7	367	2	US-08-553-058C-4 Sequence 4, App1
25	219	11.7	367	2	US-08-514-451A-4 Sequence 4, App1
26	219	11.7	367	3	US-09-170-331-4 Sequence 4, App1
27	219	11.7	367	3	US-09-510-473-4 Sequence 4, App1

28	219	11.7	367	4	US-09-048-916B-4 Sequence 4, App1
29	219	11.7	384	3	US-09-071-434-3 Sequence 3, App1
30	218	11.6	378	4	US-09-045-583-5 Sequence 5, App1
31	218	11.6	378	4	US-09-534-185-5 Sequence 5, App1
32	217	11.6	355	3	US-09-045-583-53 Sequence 53, App1
33	217	11.6	355	4	US-09-534-185-53 Sequence 53, App1
34	217	11.6	367	2	US-08-454-549-2 Sequence 2, App1
35	217	11.6	367	3	US-08-454-552-2 Sequence 2, App1
36	217	11.6	367	3	US-08-147-592A-6 Sequence 6, App1
37	217	11.6	367	3	US-08-889-108-17 Sequence 17, App1
38	217	11.6	367	3	US-08-292-694A-6 Sequence 6, App1
39	217	11.6	367	4	US-09-743-871B-15 Sequence 15, App1
40	217	11.6	367	4	US-08-986-209A-2 Sequence 2, App1
41	217	11.6	367	5	PCT-US94-10358-17 Sequence 2, App1
42	215.5	11.5	367	4	US-08-405-271A-23 Sequence 23, App1
43	215	11.5	367	4	US-08-676-351-2 Sequence 2, App1
44	213.5	11.4	380	3	US-08-676-351-5 Sequence 5, App1
45	213	11.4	337	1	US-08-153-848-46 Sequence 46, App1

ALIGNMENTS

```
RESULT 1
US-07-629-1041-3
; Sequence 3, Application US/076291041
; Patent No. 5288621
; GENERAL INFORMATION:
; APPLICANT: Gershengorn, Marvin C
; APPLICANT: Straub, Richard E
; TITLE OF INVENTION: PITUITARY TRH RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/629,1041
; FILING DATE: 19901214
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: CRF D - 995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ. ID NO. 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-629-1041-3

Query Match      12.9% Score 241; DB I; Length 393;
Best local Similarity 22.1%; Pred. No. 1.3e-13;
Matches 81; Conservative 75; Mismatches 123; Indels 88; Gaps 15;

Oy      19 PSAGCLG-VPVTVYSLLC-LGIPANILTYIIISQVARKKSYN-YLLAALADLT 75
      16 PAAVALEVQVTVITLIVITIGIVGNIMVAVNMR--TKMRPTNYLVLSAVADLM 73
Oy      76 VLFVIVFDLLEDFITNMQPOVPDKI-----TEVLEFSSIHSTIWTVPL 122
```

```

Db 74 VL-----VAGLPNITDSYGSWYGVGLCTTYLQYIGINASSCSITAF 119
QY 123 TIDRYAVCHPLKHTVSPARTKRVISVYITCFLTSIPYWM-----PNIMT-EDYIS 176
Db 120 TIERIALCHPTKAOFLCTPSRAKIIIFVMA---FTSYGLMFFLLDLNISTKANV 176
QY 177 TSVHVLIMHCFYVLPVCSIFFLINSIIV-----IYV-----208
Db 177 VSCGKISNYSPIYMDFGVYVPMILATVLYGFARILFLNIPSPDKENSKWMKN 236
QY 209 -KLARKSNFRLRG-----YSTGKTALIFTITSIFATLMAP-RIIMILYLYGAPI 257
Db 237 DSIHONKMLNATNRCNSTVSSRKQVTKMLAVVILFALMLMPYRTLVLVNSFLSSPF 296
QY 258 QNRMLVHMSDIANMLALNTAINPFLYCFISKPRRTMAATLKAFFCKQKQPVQYTNH 317
Db 297 QENMFLL---FCRICIYLSAINPVIYINLMSQKR---AARFLCNCKQKPTERKANY 348
QY 318 NFSITSS 324
Db 349 SVALNYS 355

```

RESULT 2

```

US-08-288-663A-1
/ Sequence 1, Application US/08288663A
/ Patent No. 5879896
/ GENERAL INFORMATION:
/ APPLICANT: HINUMA, Shuji
/ APPLICANT: HOSOGA, Masaki
/ APPLICANT: ONDA, Haruo
/ TITLE OF INVENTION: HUMAN TRP RECEPTOR, ITS PRODUCTION
/ NUMBER OF SEQUENCES: 15
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
/ STREET: 130 Water Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION NUMBER: US/08/288,663A
/ FILING DATE: 09-AUG-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 198309/1993
/ FILING DATE: 10-AUG-1993
/ APPLICATION NUMBER: 286986/1993
/ FILING DATE: 16-NOV-1993
/ APPLICATION NUMBER: 325215/1993
/ FILING DATE: 22-DEC-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Resnick, David S
/ REGISTRATION NUMBER: 34,235
/ REFERENCE/DOCKET NUMBER: 44612
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 398 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-288-663A-1

```

```

Query Match 12.7%; Score 238.5; DB 2; Length 398;
Best Local Similarity 22.4%; Pred. No. 2.2e-13;
Matches 80; Conservative 70; Mismatches 120; Indels 87; Gaps 13;
QY 28 VPVVYSLLLC-LGLPANILTVIILSQLVARQKSSYN-YLLAALADILVFFIVPDF 85
Db 26 VVITILVLIICGLIGINIMVLLVMR--TKMRPTNCYLSLAVADIMVL-----75
QY 86 LLEDFLINQMOPVDKI-----LEVEFSSIHSTWITVPLTIDRYAVCH 132
Db 76 -----VAGLPNITDSYGSWYGVGLCTTYLQYIGINASSCSITAFIERIALCH 129
QY 133 PLKHTVSPARTKRVISVYITCFLTSIPYWM-----PNIMT-EDYISVHVLIMT 186
Db 130 PTKAOFLCTPSRAKIIIFVMA---FTSLYGLMFFLLDLNISTKANVISCYKSRN 186
QY 187 HCFYVLPVCSIFFLINSIIVKL-----RRKSN--215
Db 187 YSPITLMDFGVYVPMILATVLYGFARILFLNIPSPDKENSKTKMDSTQNTNIN 246
QY 216 -----FRLGYSTGKTALIFTITSIFATLMAP-RIIMILYLYGAPIQNRMLVHMS 267
Db 247 VNTSNRCNSTVSSRKQVTKMLAVVILFALMLMPYRTLVLVNSFLSSPFQENMFLL---303
QY 268 DIANMLALNTAINPFLYCFISKPRRTMAATLKAFFCKQKQPVQYTNHFSITSS 324
Db 304 -FCRICIYLSAINPVIYINLMSQKR---AARFLCNCKQKPTERKANSVALNYS 355

```

RESULT 3

```

US-09-545-944-2
/ Sequence 2, Application US/09545944
/ Patent No. 6461836
/ GENERAL INFORMATION:
/ APPLICANT: AMES, ROBERT
/ APPLICANT: ELSHOUBAGY, NABIL
/ APPLICANT: MICHALOVICH, DAVID
/ APPLICANT: SARAU, HENRY
/ APPLICANT: SHABON, USMAN
/ APPLICANT: VAWTER, LISA
/ TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM RECEPTOR
/ FILE REFERENCE: GP70657-1
/ CURRENT APPLICATION NUMBER: US/09/545,944
/ PRIOR APPLICATION NUMBER: US/00-04-10
/ PRIOR FILING DATE: 1999-11-05
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 415
/ TYPE: PRT
/ ORGANISM: HOMO SAPIENS
US-09-545-944-2

```

Query Match 12.4%; Score 232; DB 4; Length 415;

Best Local Similarity 22.4%; Pred. No. 8.9e-13;

Matches 79; Conservative 77; Mismatches 112; Indels 84; Gaps 15;

```

QY 27 FVPV-VYISLLCLGLPANILTVIILSQLVARQKSSYVLLAALADILVFFIVPDF 85
Db 45 FLPVSVVYPIFVGVIGVNLVCLVILQKMKPTNY-YLFSLAVSDLVV-----LLGM 99
QY 86 LLEDFLINQMOP-----QVPDKIIEVLESSHISHTIMITVPLTIDRYAVCHPLKX 137
Db 100 PLEVYEMRNYPFLFGPVGCYFKTALFETVCFASI--LSITT-VSYERYVAIILHPRAK 155
QY 138 TVSPARTKRVISVYITCFLTSIP-----YVMWN-----IMTBDY 174
Db 156 LQSTRRLRLRIGIWMGFSVLFSLPNISIHGIXHFFNGSLVGSATCYIKPMW-----211
QY 175 ISTSVHVLIMHCFYVLPVCSIFFLINSIIVKLRR-----KSNFRLGSGTSGKT 227

```

Db 212 ----INPIIQTSPLEFLLPMTVISVLYMALRLKOKSLADEGNANIQPCRSKN 267
QY 228 AIIPTITSIPATLMAPIIMILVLYGAPIQNRW-----LVHIMSDIANMLALNT 278
Db 268 KMLFVLYLVFALCMAFPHIDLFPFSE-----VEEMTESLAVALFVHVSQV---LFYLS 320
QY 279 AINPFLYCISKRFRTMAAATLCAFPKCKQKQPVQFTNNHFTLSGPMISPA 330
Db 321 ANPPIIYNLSRRFQAAPQNVISFHK-----QMSQHD-----PQLEPA 360

RESULT 4
US-08-012-988A-2
Sequence 2, Application US/08012988A
Patent No. 5652133
GENERAL INFORMATION:
APPLICANT: Murphy, Philip M.
TITLE OF INVENTION: Cloning and Expression of Human
TITLE OF INVENTION: Macrophage Inflammatory Protein-1 alpha (MIP-1
TITLE OF INVENTION: alpha)/RANTES Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/012,988A
FILING DATE: 19930128
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-118
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-012-988A-2

Query Match 12.4%; Score 231.5; DB 1; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 ACGLGFPVNVYSSLLCLGIPANILTVIISQVARRQSSNYLLAALADIVLFFV 81
Db 31 ACGAQLLPPL-YSLVFGVIGLVGNILVAVQYKRLKNTSI-YLNLAIISDLLFLFTLP 88
QY 82 F-VDF-LLEDFLIANNQMPQVPRDKIEVLSESSIHSTIMITVPLTIDRYAVCH---PLKX 136
Db 89 FVIDYKLQDWMVGDAM---CKILSGFYTGYSIEIFITLITIDRYALVAHVPALRA 144
QY 137 HTVSFARTKRVIVSVYITCFITSIIP-YVWMPNIMTEDYISTSVNH-----181
Db 145 RVTTEGVITSIIMALAI---LASMPGLVFSKTOHEFTHTGSLHPRHSLREMKLFOAL 201
QY 182 -----VLIIHCFVTVYLVPCSIPTLNSIIVYKLRKSNFRLRGYSTGKTTAII 231
Db 202 KNLFGVLVPLVIMIICTY-----GIILRLRPNE-----KSKAVRLIF 242
QY 232 TITSIPATLMAPIIMILVLYGAPI-----QNRMLVHIMSDIANMLALNTAIVPFLY 285

Db 243 VMIIPLFWTPRYNLTILISVFODFLTHCEQSRHL-DLAVQVTEVIAYTHCQNVPIY 301
QY 286 CFISKRFR 293
Db 302 AFVGERFR 309

RESULT 5
US-08-450-393A-5
Sequence 5, Application US/08450393A
Patent No. 5707815
GENERAL INFORMATION:
APPLICANT: Charo, Israel
TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
TITLE OF INVENTION: PROTEIN RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
STREET: 5 Palo Alto Square
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94306-2155
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,393A
FILING DATE: May 25, 1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ceer, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: UCAL-237/02US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-843-5165
TELEFAX: 415-8857-0663
TELEX: 380816COOLEYPA
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-450-393A-5

Query Match 12.4%; Score 231.5; DB 1; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 ACGLGFPVNVYSSLLCLGIPANILTVIISQVARRQSSNYLLAALADIVLFFV 81
Db 31 ACGAQLLPPL-YSLVFGVIGLVGNILVAVQYKRLKNTSI-YLNLAIISDLLFLFTLP 88
QY 82 F-VDF-LLEDFLIANNQMPQVPRDKIEVLSESSIHSTIMITVPLTIDRYAVCH---PLKX 136
Db 89 FVIDYKLQDWMVGDAM---CKILSGFYTGYSIEIFITLITIDRYALVAHVPALRA 144
QY 137 HTVSFARTKRVIVSVYITCFITSIIP-YVWMPNIMTEDYISTSVNH-----181
Db 145 RVTTEGVITSIIMALAI---LASMPGLVFSKTOHEFTHTGSLHPRHSLREMKLFOAL 201
QY 182 -----VLIIHCFVTVYLVPCSIPTLNSIIVYKLRKSNFRLRGYSTGKTTAII 231
Db 202 KNLFGVLVPLVIMIICTY-----GIILRLRPNE-----KSKAVRLIF 242
QY 232 TITSIPATLMAPIIMILVLYGAPI-----QNRMLVHIMSDIANMLALNTAIVPFLY 285

Db 243 VIMIFFLFWTPYNTLILSVFODFLFTHCEQSRHL-DIAVOYTEVIAVTHCCVNPVY 301
QY 286 CFISKRFR 293
Db 302 AFVGERFR 309

RESULT 6

US-08-446-669-5
Sequence 5, Application US/08446669
Patent No. 6132987
GENERAL INFORMATION:
APPLICANT: Charo, Israel
APPLICANT: Coughlin, Shaun
TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
TITLE OF INVENTION: PROTEIN RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Cooley Godward Castro Huddleston & Tatum
STREET: 5 Palo Alto Square
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94306-2155
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,669
FILING DATE: May 25, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Neeley, Richard
REGISTRATION NUMBER: 30,092
REFERENCE/DOCKET NUMBER: UCAL-237/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-843-5000
TELEFAX: 415-857-0663
TELEX: 380816COOLEYPA
INFORMATION FOR SEO ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-446-669-5

Query Match 12.4%; Score 231.5; DB 3; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 AGGLGFVPPVYVSYLLCLGPANITVITLSQVARRKSSVNYLLAADAIVLFTIV 81
Db 31 AFGAQLLEPL-YSLVFGVIGLVNIIVLVLYQYKRLKMTSI-YLLNLAISDLFLFTLP 88
QY 82 F-VDF-LLEDFLNNQMPQVDPDKIEVLEFSSIHSTWITVPLTIDRYIACH--PLKY 136
Db 89 FVIDYKLDMDVFGDAM-----CKILSGFYTGIVSEIFILLITIDRYIAVHAFALRA 144
QY 137 HTVSYPARTKVIYVYITCFLTSIP-YWMPNWTEDYISTSVH-----181
Db 145 RTVTFGVITSIIMALAI--LASMPGLYFSKTQWEPFTHHCSLHPHESLRKWLFOAL 201
QY 182 -----VLIMHCFYVLPVPSIFLINSIIVYKLRKSNFRRLRGYSTGKTALIF 231
Db 202 KLNLFGLVLPVLMWITCYT-----GIKILRRPNE-----KSKAVRLIF 242
QY 232 TITSIFATWAPRIIMILYHLYGAPI-----QNRWLVIHMSDIANMLALNTAINFLY 285

Db 243 VIMIFFLFWTPYNTLILSVFODFLFTHCEQSRHL-DIAVOYTEVIAVTHCCVNPVY 301
QY 286 CFISKRFR 293
Db 302 AFVGERFR 309

RESULT 7

US-09-239-938-1
Sequence 1, Application US/09239938
Patent No. 6329510
GENERAL INFORMATION:
APPLICANT: Qin, Shixin
APPLICANT: Newman, Walter
APPLICANT: Kaasam, Nasim
APPLICANT: Leukosite, Inc.
TITLE OF INVENTION: ANTI-CCR1 ANTIBODIES AND METHODS OF USE
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: LKS97-13
CURRENT APPLICATION NUMBER: US/09/239,938
CURRENT FILING DATE: 1999-01-29
NUMBER OF SEO ID NOS: 1
SOFTWARE: FaestSO for Windows Version 3.0
SEO ID NO 1
LENGTH: 355
TYPE: PRT
ORGANISM: Homo sapien
US-09-239-938-1

Query Match 12.4%; Score 231.5; DB 3; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 AGGLGFVPPVYVSYLLCLGPANITVITLSQVARRKSSVNYLLAADAIVLFTIV 81
Db 31 AFGAQLLEPL-YSLVFGVIGLVNIIVLVLYQYKRLKMTSI-YLLNLAISDLFLFTLP 88
QY 82 F-VDF-LLEDFLNNQMPQVDPDKIEVLEFSSIHSTWITVPLTIDRYIACH--PLKY 136
Db 89 FVIDYKLDMDVFGDAM-----CKILSGFYTGIVSEIFILLITIDRYIAVHAFALRA 144
QY 137 HTVSYPARTKVIYVYITCFLTSIP-YWMPNWTEDYISTSVH-----181
Db 145 RTVTFGVITSIIMALAI--LASMPGLYFSKTQWEPFTHHCSLHPHESLRKWLFOAL 201
QY 182 -----VLIMHCFYVLPVPSIFLINSIIVYKLRKSNFRRLRGYSTGKTALIF 231
Db 202 KLNLFGLVLPVLMWITCYT-----GIKILRRPNE-----KSKAVRLIF 242
QY 232 TITSIFATWAPRIIMILYHLYGAPI-----QNRWLVIHMSDIANMLALNTAINFLY 285
Db 243 VIMIFFLFWTPYNTLILSVFODFLFTHCEQSRHL-DIAVOYTEVIAVTHCCVNPVY 301
QY 286 CFISKRFR 293
Db 302 AFVGERFR 309

RESULT 8

US-09-886-319A-14
Sequence 14, Application US/09886319A
Patent No. 6586185
GENERAL INFORMATION:
APPLICANT: Wolf, Eckard
APPLICANT: Werner, Sabine
APPLICANT: Halle, Jörn-Peter
APPLICANT: Regenbogen, Johannes
APPLICANT: Goppelt, Andreas
TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
TITLE OF INVENTION: The Diagnosis or Treatment of Skin Disorders and Wound
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
Active Substances

QY 232 TTSIFATLAPRIIMILYHLYGAP-----QNRVLVHMSDIANMLALNTAINPFLY 285
DB 243 VIMIFFLEFWTPYNTLILSVFQDFLFTHECQSRHL-DLAVQVEVIAVYTHCCVNPVY 301
QY 286 CFIKRRFR 293
DB 302 AFVGERFR 309

RESULT 11

US-09-625-573-5
Sequence 5, Application US/09625573
Patent No. 6730301
GENERAL INFORMATION:
APPLICANT: Charo, Israel
Coughlin, Shaun
TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOTACTANT
PROTEIN RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Godward Castro Huddleston & Tatum
STREET: 5 Palo Alto Square
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94306-2155
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/625,573
FILING DATE: 25-Jul-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/446,669
FILING DATE: May 25, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Neeley, Richard
REGISTRATION NUMBER: 30,092
REFERENCE/DOCKET NUMBER: UCAL-237/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-843-5000
TELEFAX: 415-857-0663
TELEX: 380816COOLEYPA
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-625-573-5

Query Match 12.4%; Score 231.5; DB 4; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;

Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 ACGAGFVPPVYVYSLCLGIPANILTVIILSQLVARRQSSVNYLLAAADIVLPFIY 81
DB 31 AFGAQLPPL-YSLVFGVIGVNIIVLVVQYKRLKMTSI-YLNLTAISDLFLFTLP 88
QY 82 F-VDP-LLEDFILNMQMPQVDPDKIEVLEFSSIHSTWITVPLTIDRYAVCH--PLKY 136
DB 89 FWIDYKLDWVFGDAM---CKILSGFYTGLYSEIFILLITIDRIALVHAVPALRA 144
QY 137 HTVSYPARTKRVISVYITCFLTSP-YVWPNWTEYISTSVHH-----181
DB 145 RTVTFGVITSIITMALAI---LASMFGLYFSKTQWFTHTTCSLHPHESLRKWLFOAL 201

QY 182 -----VLWICFTVYVPCSIFFILNSIIVYKLRKSNRRLRGSTGTAILE 231
DB 202 KNLFGVLVPLVMICYT-----GIKILRRPNE-----KSKAVRLIF 242
QY 232 TTSIFATLAPRIIMILYHLYGAP-----QNRVLVHMSDIANMLALNTAINPFLY 285
DB 243 VIMIFFLEFWTPYNTLILSVFQDFLFTHECQSRHL-DLAVQVEVIAVYTHCCVNPVY 301
QY 286 CFIKRRFR 293
DB 302 AFVGERFR 309

RESULT 12

US-09-960-547-1
Sequence 1, Application US/09960547
Patent No. 6756035
GENERAL INFORMATION:
APPLICANT: Qin, Shixin
APPLICANT: Newman, Walter
APPLICANT: Kassam, Nasim
TITLE OF INVENTION: ANTI-CCR1 ANTIBODIES AND METHODS OF USE
FILE REFERENCE: 1855.1048-010
CURRENT APPLICATION NUMBER: US/09/960,547
PRIOR FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: US 09/239,938
NUMBER OF SEQ ID NOS: 1
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 1
LENGTH: 355
TYPE: PRT
ORGANISM: Homo sapiens
US-09-960-547-1

Query Match 12.4%; Score 231.5; DB 4; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;

Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 ACGAGFVPPVYVYSLCLGIPANILTVIILSQLVARRQSSVNYLLAAADIVLPFIY 81
DB 31 AFGAQLPPL-YSLVFGVIGVNIIVLVVQYKRLKMTSI-YLNLTAISDLFLFTLP 88
QY 82 F-VDP-LLEDFILNMQMPQVDPDKIEVLEFSSIHSTWITVPLTIDRYAVCH--PLKY 136
DB 89 FWIDYKLDWVFGDAM---CKILSGFYTGLYSEIFILLITIDRIALVHAVPALRA 144
QY 137 HTVSYPARTKRVISVYITCFLTSP-YVWPNWTEYISTSVHH-----181
DB 145 RTVTFGVITSIITMALAI---LASMFGLYFSKTQWFTHTTCSLHPHESLRKWLFOAL 201
QY 182 -----VLWICFTVYVPCSIFFILNSIIVYKLRKSNRRLRGSTGTAILE 231
DB 202 KNLFGVLVPLVMICYT-----GIKILRRPNE-----KSKAVRLIF 242
QY 232 TTSIFATLAPRIIMILYHLYGAP-----QNRVLVHMSDIANMLALNTAINPFLY 285
DB 243 VIMIFFLEFWTPYNTLILSVFQDFLFTHECQSRHL-DLAVQVEVIAVYTHCCVNPVY 301
QY 286 CFIKRRFR 293
DB 302 AFVGERFR 309

RESULT 13
PCT-US95-00476-5
Sequence 5, Application PC/TUS9500476
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOTACTANT
PROTEIN RECEPTORS
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Robbins, Berlinger & Carson
 STREET: 201 N. Figueroa Street, 5th floor
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90012-2628
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/00476
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Berlinger, Robert
 REGISTRATION NUMBER: 20,121
 REFERENCE/DOCKET NUMBER: 5555-291
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 310-977-1001
 TELEFAX: 310-977-1003
 TELEX:
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 355 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 PCT-US95-00476-5

Query Match 12.4%; Score 231.5; DB 5; Length 355;
 Best local Similarity 26.3%; Fred. No. 8.2e-13;
 Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 ACGLGFVYVYVYSLICLGPANILVTIISQVARRKSSSYVLLAALADLVLPFY 81
 DB 31 AFGAQLLPFLVSLVFIYVIGVNLVVLVQYKRLKMTSI-YLNLALSDLPFLFY 88
 QY 82 F-VDF-LLEDPLINMGPVDPDKIIEVLEFSIHISITWIVPLTIDRYAVCH---FLKX 136
 DB 89 FWDYLDKDDWGDWDM---CCLSGFYTGSLYSEIFLITLIDRYALVAVPALRA 144
 QY 137 HTVSFARTRKVIYSVYITCFLTSIP-YVWPMIWTEDYSTSVH-----181
 DB 145 RVTTPGVYISIIIMALAI---LASMFGVYSKTOMETHHTCSLHPHESLRMKLPQAL 201
 QY 182 -----VLIWHCFVYLVPCSIFFLINSIIYVKLRKSNFRLRGYSTGKTAALIF 231
 DB 202 KNLPLGLVLPFLVMIICYT-----GIKILRRPNE-----KSKAVRLIF 242
 QY 232 TTTSIATLMAPIIMILVLYGAFI-----QNRMLVHIMSDIANMLALNAINPFLY 285
 DB 243 VMIIFPLFPTPNLITLISVFOPLFTHCEQSRHL-DLAVGVTEVIAYTHCCVAPVIY 301
 QY 286 CFISKRFR 293
 DB 302 APVGERFR 309

RESULT 14
 US-08-288-663A-15
 Sequence 15, Application US/08288663A
 Patent No. 5879896
 GENERAL INFORMATION:
 APPLICANT: HINDA, Shuji
 APPLICANT: ONDA, Haruo
 APPLICANT: HOSODA, Masaki
 TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION
 TITLE OF INVENTION: AND USE

NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DIXE, BRONSTEIN, ROBERTS & CUSMAN, LLP
 STREET: 130 Water Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: PasteSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/288,663A
 FILING DATE: 09-AUG-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 198309/1993
 FILING DATE: 10-AUG-1993
 APPLICATION NUMBER: 286986/1993
 FILING DATE: 16-NOV-1993
 APPLICATION NUMBER: 325215/1993
 FILING DATE: 22-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Resnick, David S
 REGISTRATION NUMBER: 34,235
 REFERENCE/DOCKET NUMBER: 44612
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-523-3400
 TELEFAX: 617-523-6440
 TELEX:
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 398 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-288-663A-15

Query Match 12.1%; Score 226.5; DB 2; Length 398;
 Best local Similarity 22.6%; Pred. No. 2.6e-12;
 Matches 83; Conservative 70; Mismatches 125; Indels 89; Gaps 15;

QY 19 FGSAGLGF-VPVYVYSLIC-LGLFANILVTIISQVARRKSSYN-YLALAALADLV 75
 DB 16 FOVAVALBEYQVVTIILVAVICIGIVGNIMVVLVWR--TKMRRTNCTVLSLAVADLM 73
 QY 76 VLPFYVDFLLEDPLINMGPVDPDKI-----IEVLEFSIHISITWIVPL 122
 DB 74 VL-----VAAGLFNITDSIYGSWYGVGCLCTIYQYLGINSKSGITAF 119
 QY 123 TIDRYAVCHPDKHTVSFARTRKVIYSVYITCFLTSIPYVW-----FNITW-EDYIS 176
 DB 120 TIERIYA-CHFIKAQGLCTFSRAKKIILFVMA---FTSYCMVLPFLDLNISTYDALV 175
 QY 177 TSVHVLVHICFTVYLVPCSIFFLINSIIYVKL-----210
 DB 176 ISGKYISNHYSPYIMDFGVYVFMILATLVGFIARILFNFIFSDPKENSKTWYN 235
 QY 211 ---RRKSNFRLG-----YSTGKTAALIFTTISIFATLMAPI-RIMILVLYGAFI 257
 DB 236 DSTHONKNNLTTNRCFNSVTSRKQVTKMLAVVILFRLMPEFRTLVVNSKLSIF 295
 QY 258 QNRMLVHIMSDIANMLALNAINPFLYCFISKRFRYMAAATLKAFKQCKQFVOPFYTH 317
 DB 296 QENWPLF-----FCRICTYLSAINPIYVYLMQOKR-----AARLCKNCQKQFTEKAANY 347
 QY 318 NFSITSS 324
 DB 348 SVALNYS 354

```
RESULT 15
US-08-833-752-9
/ Sequence 9, Application US/08833752
/ Patent No. 6448375
/ GENERAL INFORMATION:
/ APPLICANT: SAMSON, MICHEL
/ APPLICANT: PARMENTIER, MARC
/ APPLICANT: VASSART, GILBERT
/ APPLICANT: LIBERT, FREDERICK
/ TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
/ NUMBER OF SEQUENCES: 17
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: Knobbé, Martens, Olsson & Bear
/ STREET: 620 Newport Center Drive 16th Floor
/ CITY: Newport Beach
/ STATE: CA
/ COUNTRY: U.S.A.
/ ZIP: 92660
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentln Release #1.0, Version #1.25 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/833,752
/ FILING DATE: 9-APR-1997
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Altman, Daniel E
/ REGISTRATION NUMBER: 34,115
/ REFERENCE/DOCKET NUMBER:
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 355 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: No. 6448375e
/ US-08-833-752-9

Query Match      11.9%; Score 223.5; DB 4; Length 355;
Best Local Similarity 26.8%; Pred.No.4.3e-12;
Matches 79; Conservative 64; Mismatches 113; Indels 39; Gaps 15;

QY      22  ACGCGFVPPVYVYSLICGIPANILTVIILSQLVAROKSYNYVLALAAADILVLFPIV 81
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      31  AFGAQLPLPL-YSLVFGVIGLVGNILVVLVLYQKRLKNTSI-YLANLAISDLPLFIPTIP 88
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      82  F-VDF-LLEDPLINQMPOVPDKIIIEVLEFSSIHISIMITPPLTDRIYAVCH---PLKY 136
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      89  FWDYKLDKDWYFGDAM---CKIISGFYTGVSIFPIILITDRIYLAIVHAVFAIRA 144
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      137 HTVSYPARTKVIIVSVYITCFLTSLIP-YYMWNINWTEDIYSTSVH--HVLV--W----- 185
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      145 RVTYFGVLTISIIMIAI---IASMGLYFSTQWETHTHTCSLHPHESLREWKLFQAL 201
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      186 -IHCFYVIVPCSIFFILNLSIVYKLRKSNFRLNGYSTGKTTAILFTTISIFATLMAPR 244
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      202 KLNLFGLVL-PLLVNMLICYTGIIKILRLRPNEK-----KSAVRLLIFVIMILFFLEWIPY 255
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      245 IIMILYHLYGAPI-----QNRWLVIHMSDIANMLALNTAINPFLYCFISKRR 293
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      256 NLTIIISVFQDPLFTHECEQSRHL-DLAVQVTEVIATYHCCVNEVIYAFVGERR 309
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
```

Search completed: December 14, 2004, 20:28:13
Job time : 17.4665 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:23:14 ; Search time 49.3994 Seconds
(without alignments)
2552.337 Million cell updates/sec

Title: US-09-995-225B-16

Sequence: 1 MHTRAHLAANSSLSWSPG.....CIKMLVQYDKKPKIKVSP 353

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1872	100.0	353	9	US-09-995-225-16
2	1872	100.0	353	10	US-09-995-225-16
3	1872	100.0	353	14	US-10-094-417-2
4	1872	100.0	353	15	US-10-296-294A-5
5	1872	100.0	353	15	US-10-450-590-2
6	1872	100.0	353	16	US-10-779-104-2
7	1868	99.8	353	10	US-09-813-433-22
8	1868	99.8	353	14	US-10-174-364-22
9	1868	99.8	353	15	US-10-246-583-22
10	1868	99.8	353	16	US-10-689-832-22
11	1857	99.2	350	10	US-09-813-433-60
12	1857	99.2	350	14	US-10-174-364-60
13	1857	99.2	350	15	US-10-246-583-60

14	1857	99.2	350	16	US-10-689-832-60	Sequence 60, Appl
15	1857	99.2	372	10	US-09-813-432-20	Sequence 20, Appl
16	1857	99.2	372	14	US-10-219-834-19	Sequence 19, Appl
17	1857	99.2	372	14	US-10-314-076-2	Sequence 2, Appl
18	1857	99.2	372	14	US-10-174-364-20	Sequence 20, Appl
19	1857	99.2	372	15	US-10-333-946-6	Sequence 6, Appl
20	1857	99.2	372	15	US-10-246-583-20	Sequence 20, Appl
21	1857	99.2	372	16	US-10-689-832-20	Sequence 20, Appl
22	1857	99.2	372	17	US-10-712-615-103	Sequence 103, App
23	1853	99.0	350	10	US-09-813-432-59	Sequence 59, Appl
24	1853	99.0	350	14	US-10-174-364-59	Sequence 59, Appl
25	1853	99.0	350	15	US-10-246-583-59	Sequence 59, Appl
26	1853	99.0	350	16	US-10-689-832-59	Sequence 59, Appl
27	1844	98.5	353	14	US-10-012-140-11	Sequence 11, Appl
28	1841	98.3	369	14	US-10-314-076-4	Sequence 4, Appl
29	1756	93.8	345	14	US-10-094-417-20	Sequence 20, Appl
30	1643	87.8	318	14	US-10-174-364-85	Sequence 85, Appl
31	1643	87.8	318	15	US-10-246-583-85	Sequence 85, Appl
32	1638	87.5	333	14	US-10-079-384-26	Sequence 26, Appl
33	1638	87.5	343	15	US-10-450-530-8	Sequence 8, Appl
34	1638	87.5	356	15	US-10-450-530-7	Sequence 7, Appl
35	1638	87.5	385	15	US-10-343-650A-22	Sequence 22, Appl
36	1471	78.6	287	10	US-09-791-932-66	Sequence 66, Appl
37	1426	76.2	272	10	US-09-813-432-57	Sequence 57, Appl
38	1426	76.2	272	14	US-10-174-364-57	Sequence 57, Appl
39	1426	76.2	272	14	US-10-246-583-57	Sequence 57, Appl
40	1426	76.2	272	15	US-10-689-832-57	Sequence 57, Appl
41	1426	76.2	272	16	US-10-689-832-58	Sequence 58, Appl
42	1426	76.2	272	16	US-10-225-567A-587	Sequence 587, App
43	1317	70.4	265	14	US-10-314-076-14	Sequence 14, Appl
44	1317	70.4	265	14	US-10-314-076-14	Sequence 14, Appl
45	912.5	48.7	313	10	US-09-791-932-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-995-225-16
; Sequence 16, Application US/09995225
; Publication No. US20020193584A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pridle, Cameron
; TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917

;; PRIOR FILING DATE: 2001-05-14
;; PRIOR APPLICATION NUMBER: 60/309,208
;; PRIOR FILING DATE: 2001-07-31
;; NUMBER OF SEQ ID NOS: 67
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 16
;; LENGTH: 353
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: No. US20020193584A1 Sequence
US-09-995-225-16

Query Match 100.0%; Score 1872; DB 9; Length 353;
Best Local Similarity 100.0%; Pred. No. 2,3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSGAGLGFVPPVYYSLLCLGLPANILTVIIISQVARROK 60
DB 1 MEHTHAHLAANSSLSWSPGSGAGLGFVPPVYYSLLCLGLPANILTVIIISQVARROK 60
QY 61 SSVNYLLAAADILVLFVIFVDFLEDFILNMQMPQVPDKIIIEVLEFSSIHSTWTV 120
DB 61 SSVNYLLAAADILVLFVIFVDFLEDFILNMQMPQVPDKIIIEVLEFSSIHSTWTV 120
QY 121 PLTIDRYIAVCHPLKHTVSPARTKRVISVYITCFLTSIPYWMNMTEDYISTSVH 180
DB 121 PLTIDRYIAVCHPLKHTVSPARTKRVISVYITCFLTSIPYWMNMTEDYISTSVH 180
QY 181 HVLIMHCFVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISRATL 240
DB 181 HVLIMHCFVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISRATL 240
QY 241 WAPRIIMLVHLYGAPIONRWLVHMSDIANMLALNTAINFLYCFISRRFRMAAATL 300
DB 241 WAPRIIMLVHLYGAPIONRWLVHMSDIANMLALNTAINFLYCFISRRFRMAAATL 300
QY 301 KAFKCKOKOPVQFTYNNFSTSSPWISPA NSHCIMLVYQYDKNGKPIKVP 353
DB 301 KAFKCKOKOPVQFTYNNFSTSSPWISPA NSHCIMLVYQYDKNGKPIKVP 353

RESULT 2

US-09-995-225-16
;; Sequence 16, Application US/09995225
;; Publication No. US20030139588A9
;; GENERAL INFORMATION:
;; APPLICANT: Chen, Ruoping
;; APPLICANT: Chu, Zhi Liang
;; APPLICANT: Dang, Huang T.
;; APPLICANT: Lowitz, Kevin P.
;; APPLICANT: Priddy, Catherine
;; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human C
;; FILE REFERENCE: AREN-0308
;; CURRENT APPLICATION NUMBER: US/09/995,225
;; PRIOR APPLICATION NUMBER: 09/170,496
;; PRIOR FILING DATE: 1998-10-13
;; PRIOR APPLICATION NUMBER: PCT/US99/23938
;; PRIOR FILING DATE: 1998-10-13
;; PRIOR APPLICATION NUMBER: 60/253,404
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/255,366
;; PRIOR FILING DATE: 2000-12-12
;; PRIOR APPLICATION NUMBER: 60/270,286
;; PRIOR FILING DATE: 2001-02-20
;; PRIOR APPLICATION NUMBER: 60/282,365
;; PRIOR FILING DATE: 2001-04-06
;; PRIOR APPLICATION NUMBER: 60/270,266
;; PRIOR FILING DATE: 2001-02-20
;; PRIOR APPLICATION NUMBER: 60/282,032
;; PRIOR FILING DATE: 2001-04-06

;; PRIOR APPLICATION NUMBER: 60/282,358
;; PRIOR FILING DATE: 2001-04-06
;; PRIOR APPLICATION NUMBER: 60/282,356
;; PRIOR FILING DATE: 2001-04-06
;; PRIOR APPLICATION NUMBER: 60/290,917
;; PRIOR FILING DATE: 2001-05-14
;; PRIOR APPLICATION NUMBER: 60/309,208
;; PRIOR FILING DATE: 2001-07-31
;; NUMBER OF SEQ ID NOS: 67
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 16
;; LENGTH: 353
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: No. US20030139588A9e1 Sequence
US-09-995-225-16

Query Match 100.0%; Score 1872; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 2,3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSGAGLGFVPPVYYSLLCLGLPANILTVIIISQVARROK 60
DB 1 MEHTHAHLAANSSLSWSPGSGAGLGFVPPVYYSLLCLGLPANILTVIIISQVARROK 60
QY 61 SSVNYLLAAADILVLFVIFVDFLEDFILNMQMPQVPDKIIIEVLEFSSIHSTWTV 120
DB 61 SSVNYLLAAADILVLFVIFVDFLEDFILNMQMPQVPDKIIIEVLEFSSIHSTWTV 120
QY 121 PLTIDRYIAVCHPLKHTVSPARTKRVISVYITCFLTSIPYWMNMTEDYISTSVH 180
DB 121 PLTIDRYIAVCHPLKHTVSPARTKRVISVYITCFLTSIPYWMNMTEDYISTSVH 180
QY 181 HVLIMHCFVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISRATL 240
DB 181 HVLIMHCFVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISRATL 240
QY 241 WAPRIIMLVHLYGAPIONRWLVHMSDIANMLALNTAINFLYCFISRRFRMAAATL 300
DB 241 WAPRIIMLVHLYGAPIONRWLVHMSDIANMLALNTAINFLYCFISRRFRMAAATL 300
QY 301 KAFKCKOKOPVQFTYNNFSTSSPWISPA NSHCIMLVYQYDKNGKPIKVP 353
DB 301 KAFKCKOKOPVQFTYNNFSTSSPWISPA NSHCIMLVYQYDKNGKPIKVP 353

RESULT 3

US-10-094-417-2
;; Sequence 2, Application US/10094417
;; Publication No. US20030045685A1
;; GENERAL INFORMATION:
;; APPLICANT: Tian, Hui
;; APPLICANT: Zhao, Jiagang
;; APPLICANT: Chen, Jin-Long
;; APPLICANT: Cutler, Gene
;; APPLICANT: Tularik Inc.
;; TITLE OF INVENTION: NO. US20030045685A1e1 Receptors
;; FILE REFERENCE: 018781-00811005
;; CURRENT APPLICATION NUMBER: US/10/094,417
;; PRIOR FILING DATE: 2002-06-10
;; PRIOR APPLICATION NUMBER: US 09/802,803
;; PRIOR FILING DATE: 2001-03-09
;; PRIOR APPLICATION NUMBER: US 60/276,649
;; PRIOR FILING DATE: 2001-03-16
;; NUMBER OF SEQ ID NOS: 43
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 353
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20

US-10-094-417-2

Query Match 100.0%; Score 1872; DB 14; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLANSLSWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIILSQLVARROK 60
DB 1 MEHTAHLANSLSWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIILSQLVARROK 60
QY 61 SSYNVLLAALAAADILVLFVFDPLLEDFILNMOMPOVDKIEVLEFSSIHITSIMITY 120
DB 61 SSYNVLLAALAAADILVLFVFDPLLEDFILNMOMPOVDKIEVLEFSSIHITSIMITY 120
QY 121 PLTIDRYIAVCHPLKHTVSYPARTRKVIYSVITCFELTSPYMPMNITWEDYISTSVAH 180
DB 121 PLTIDRYIAVCHPLKHTVSYPARTRKVIYSVITCFELTSPYMPMNITWEDYISTSVAH 180
QY 181 HVLIMHCFVTVLVPSCIFPILNSIIYVYKLRRKSNRNLGSGTKTALFTTISIFATL 240
DB 181 HVLIMHCFVTVLVPSCIFPILNSIIYVYKLRRKSNRNLGSGTKTALFTTISIFATL 240
QY 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300
DB 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300
QY 301 KAFPKQKOPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVPSP 353
DB 301 KAFPKQKOPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVPSP 353

RESULT 4
US-10-296-294A-5

/ Sequence 5, Application US/10296294A
/ Publication No. US20040029224A1
/ GENERAL INFORMATION:
/ APPLICANT: TERAQ, Yasuko
/ APPLICANT: MATSUI, Hideki
/ APPLICANT: SHINJANI, Yasuaki
/ TITLE OF INVENTION: No. US20040029224A1 G Protein-Coupled Receptor and its DNA
/ FILE REFERENCE: 2734 USOP
/ CURRENT APPLICATION NUMBER: US/10/296, 294A
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: PCT/JP01/04643
/ PRIOR FILING DATE: 2001-06-01
/ PRIOR APPLICATION NUMBER: JP 2000-170446
/ PRIOR FILING DATE: 2000-06-02
/ PRIOR APPLICATION NUMBER: JP 2000-194926
/ NUMBER OF SEQ ID NOS: 8
/ SEQ ID NO 5
/ LENGTH: 353
/ TYPE: PRT
/ ORGANISM: Human
US-10-296-294A-5

Query Match 100.0%; Score 1872; DB 15; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLANSLSWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIILSQLVARROK 60
DB 1 MEHTAHLANSLSWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIILSQLVARROK 60
QY 61 SSYNVLLAALAAADILVLFVFDPLLEDFILNMOMPOVDKIEVLEFSSIHITSIMITY 120
DB 61 SSYNVLLAALAAADILVLFVFDPLLEDFILNMOMPOVDKIEVLEFSSIHITSIMITY 120
QY 121 PLTIDRYIAVCHPLKHTVSYPARTRKVIYSVITCFELTSPYMPMNITWEDYISTSVAH 180
DB 121 PLTIDRYIAVCHPLKHTVSYPARTRKVIYSVITCFELTSPYMPMNITWEDYISTSVAH 180
QY 181 HVLIMHCFVTVLVPSCIFPILNSIIYVYKLRRKSNRNLGSGTKTALFTTISIFATL 240

DB 181 HVLIMHCFVTVLVPSCIFPILNSIIYVYKLRRKSNRNLGSGTKTALFTTISIFATL 240

QY 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300
DB 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300

QY 301 KAFPKQKOPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVPSP 353
DB 301 KAFPKQKOPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVPSP 353

RESULT 5
US-10-450-590-2

/ Sequence 2, Application US/10450590
/ Publication No. US20040076985A1
/ GENERAL INFORMATION:
/ APPLICANT: Bayer AG
/ TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
/ FILE REFERENCE: LIO316 Foreign Countries
/ CURRENT APPLICATION NUMBER: US/10/450,590
/ PRIOR FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: US 60/254,923
/ PRIOR FILING DATE: 2000-12-14
/ PRIOR APPLICATION NUMBER: US 60/280,110
/ PRIOR FILING DATE: 2001-04-02
/ PRIOR APPLICATION NUMBER: US 60/299,474
/ PRIOR FILING DATE: 2001-06-21
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 2
/ LENGTH: 353
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-450-590-2

Query Match 100.0%; Score 1872; DB 15; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLANSLSWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIILSQLVARROK 60
DB 1 MEHTAHLANSLSWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIILSQLVARROK 60
QY 61 SSYNVLLAALAAADILVLFVFDPLLEDFILNMOMPOVDKIEVLEFSSIHITSIMITY 120
DB 61 SSYNVLLAALAAADILVLFVFDPLLEDFILNMOMPOVDKIEVLEFSSIHITSIMITY 120
QY 121 PLTIDRYIAVCHPLKHTVSYPARTRKVIYSVITCFELTSPYMPMNITWEDYISTSVAH 180
DB 121 PLTIDRYIAVCHPLKHTVSYPARTRKVIYSVITCFELTSPYMPMNITWEDYISTSVAH 180
QY 181 HVLIMHCFVTVLVPSCIFPILNSIIYVYKLRRKSNRNLGSGTKTALFTTISIFATL 240
DB 181 HVLIMHCFVTVLVPSCIFPILNSIIYVYKLRRKSNRNLGSGTKTALFTTISIFATL 240
QY 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300
DB 241 MAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300
QY 301 KAFPKQKOPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVPSP 353
DB 301 KAFPKQKOPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVPSP 353

RESULT 6
US-10-779-104-2

/ Sequence 2, Application US/10779104
/ Publication No. US20040161799A1
/ GENERAL INFORMATION:
/ APPLICANT: Andrew J. Murphy
/ APPLICANT: Susan Croli-Kalish
/ TITLE OF INVENTION: KOR3L-like Proteins and Methods of Modulating KOR3L-Mediated Activity

FILE REFERENCE: REG 1000A
CURRENT APPLICATION NUMBER: US/10/779,104
CURRENT FILING DATE: 2004-02-13
PRIOR APPLICATION NUMBER: 60/447,447
PRIOR FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: 60/495,577
PRIOR FILING DATE: 2003-08-14
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 2
LENGTH: 353
TYPE: PRT
ORGANISM: homo sapiens
US-10-779-104-2

Query Match 100.0%; Score 1872; DB 16; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSSLSWSPGACGFGFVYVYSLCLGLPANILTVIILISQVARRK 60
DB 1 MEHTAHLAANSSLSWSPGACGFGFVYVYSLCLGLPANILTVIILISQVARRK 60
QY 61 SSYNVLAALAAADILVFFIVFDLLEDFILNMQMPQVDPKIEVLEFSSIHSTIWTV 120
DB 61 SSYNVLAALAAADILVFFIVFDLLEDFILNMQMPQVDPKIEVLEFSSIHSTIWTV 120
QY 121 PLTIDRYAVCHPDKHTVSPARTKRVISVYITCTLSIPYWMENWTEDEYISTSVH 180
DB 121 PLTIDRYAVCHPDKHTVSPARTKRVISVYITCTLSIPYWMENWTEDEYISTSVH 180
QY 181 HVLWIMHCFYVYVPCSIFFILNSIIVYKLRKSNFRLRGSTGKTTALIFTTISFATL 240
DB 181 HVLWIMHCFYVYVPCSIFFILNSIIVYKLRKSNFRLRGSTGKTTALIFTTISFATL 240
QY 241 WAPRIIMLVHLYGAPLQNRMLVHMSDIANMLALNTAINFLYCFISKRFTMAAATL 300
DB 241 WAPRIIMLVHLYGAPLQNRMLVHMSDIANMLALNTAINFLYCFISKRFTMAAATL 300
QY 301 KAPFKCKQKQPVQFYTNHNFISITSSPWI SPANSHCIKMLVYQYDKNGKPIKVP 353
DB 301 KAPFKCKQKQPVQFYTNHNFISITSSPWI SPANSHCIKMLVYQYDKNGKPIKVP 353

RESULT 7

US-09-813-432-22
Sequence 22, Application US/09813432
Publication No. US20030148465A1
GENERAL INFORMATION:
APPLICANT: Taupier Jr., Raymond J
APPLICANT: Majmuder, Kamud
APPLICANT: Spaderna, Steven K
APPLICANT: Smithson, Glenda
APPLICANT: Mezes, Peter S
APPLICANT: Verneet, Corine A. M.
TITLE OF INVENTION: No. US20030148465A1el Polypeptides and Amino Acids Encoding Same
FILE REFERENCE: 15966-729
CURRENT APPLICATION NUMBER: US/09/813,432
CURRENT FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28

PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,836
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/193,843
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 22
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
US-09-813-432-22

Query Match 99.8%; Score 1868; DB 10; Length 353;
Best Local Similarity 99.7%; Pred. No. 5.3e-163;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSSLSWSPGACGFGFVYVYSLCLGLPANILTVIILISQVARRK 60
DB 1 MEHTAHLAANSSLSWSPGACGFGFVYVYSLCLGLPANILTVIILISQVARRK 60
QY 61 SSYNVLAALAAADILVFFIVFDLLEDFILNMQMPQVDPKIEVLEFSSIHSTIWTV 120
DB 61 SSYNVLAALAAADILVFFIVFDLLEDFILNMQMPQVDPKIEVLEFSSIHSTIWTV 120
QY 121 PLTIDRYAVCHPDKHTVSPARTKRVISVYITCTLSIPYWMENWTEDEYISTSVH 180
DB 121 PLTIDRYAVCHPDKHTVSPARTKRVISVYITCTLSIPYWMENWTEDEYISTSVH 180
QY 181 HVLWIMHCFYVYVPCSIFFILNSIIVYKLRKSNFRLRGSTGKTTALIFTTISFATL 240
DB 181 HVLWIMHCFYVYVPCSIFFILNSIIVYKLRKSNFRLRGSTGKTTALIFTTISFATL 240
QY 241 WAPRIIMLVHLYGAPLQNRMLVHMSDIANMLALNTAINFLYCFISKRFTMAAATL 300
DB 241 WAPRIIMLVHLYGAPLQNRMLVHMSDIANMLALNTAINFLYCFISKRFTMAAATL 300
QY 301 KAPFKCKQKQPVQFYTNHNFISITSSPWI SPANSHCIKMLVYQYDKNGKPIKVP 353
DB 301 KAPFKCKQKQPVQFYTNHNFISITSSPWI SPANSHCIKMLVYQYDKNGKPIKVP 353

RESULT 8

US-10-174-364-22
Sequence 22, Application US/10174364
Publication No. US20030216308A1
GENERAL INFORMATION:
APPLICANT: Anderson et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-729CIP2
CURRENT APPLICATION NUMBER: US/10/174,364
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28

```

; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-364-22

Query Match          99.8%; Score 1868; DB 14; Length 353;
Best Local Similarity 99.7%; Pred. No. 5,3e-163;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MEHTHAHLAANSSLSWMSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVAROK 60
        1 MEHTHAHLAANSSLSWMSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVAROK 60
DB      1 MEHTHAHLAANSSLSWMSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVAROK 60
QY      61 SSNNYLALAAADILVLPFIVFDLLEDFILNMQPQVDPDKIEVLEFSSHTSIWTV 120
        61 SSNNYLALAAADILVLPFIVFDLLEDFILNMQPQVDPDKIEVLEFSSHTSIWTV 120
DB      61 SSNNYLALAAADILVLPFIVFDLLEDFILNMQPQVDPDKIEVLEFSSHTSIWTV 120
QY      121 PLTIDRYIACHPDLKHTVSYPARTRKIVSVYITCFELISIPYMMPNITWEDYISTSVH 180
        121 PLTIDRYIACHPDLKHTVSYPARTRKIVSVYITCFELISIPYMMPNITWEDYISTSVH 180
DB      121 PLTIDRYIACHPDLKHTVSYPARTRKIVSVYITCFELISIPYMMPNITWEDYISTSVH 180
QY      181 HVLWIMHCFVTVLPSCIFPILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
        181 HVLWIMHCFVTVLPSCIFPILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
DB      181 HVLWIMHCFVTVLPSCIFPILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
QY      241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINPFLYCFISGRFTMAAATL 300
        241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINPFLYCFISGRFTMAAATL 300
DB      241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINPFLYCFISGRFTMAAATL 300
QY      301 KAFKCKQKQPVQPYTHNFSITSSPWPISPANSHCICMLVYQYDKNGKPIKVP 353
        301 KAFKCKQKQPVQPYTHNFSITSSPWPISPANSHCICMLVYQYDKNGKPIKVP 353
DB      301 KAFKCKQKQPVQPYTHNFSITSSPWPISPANSHCICMLVYQYDKNGKPIKVP 353

RESULT 9
US-10-246-583-22
; Sequence 22, Application US/10246583
; Publication No. US20040058862A1
; GENERAL INFORMATION:
; APPLICANT: Majumder
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
```

```

; SEQ ID NO 22
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-246-583-22

Query Match          99.8%; Score 1868; DB 15; Length 353;
Best Local Similarity 99.7%; Pred. No. 5,3e-163;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MEHTHAHLAANSSLSWMSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVAROK 60
        1 MEHTHAHLAANSSLSWMSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVAROK 60
DB      1 MEHTHAHLAANSSLSWMSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVAROK 60
QY      61 SSNNYLALAAADILVLPFIVFDLLEDFILNMQPQVDPDKIEVLEFSSHTSIWTV 120
        61 SSNNYLALAAADILVLPFIVFDLLEDFILNMQPQVDPDKIEVLEFSSHTSIWTV 120
DB      61 SSNNYLALAAADILVLPFIVFDLLEDFILNMQPQVDPDKIEVLEFSSHTSIWTV 120
QY      121 PLTIDRYIACHPDLKHTVSYPARTRKIVSVYITCFELISIPYMMPNITWEDYISTSVH 180
        121 PLTIDRYIACHPDLKHTVSYPARTRKIVSVYITCFELISIPYMMPNITWEDYISTSVH 180
DB      121 PLTIDRYIACHPDLKHTVSYPARTRKIVSVYITCFELISIPYMMPNITWEDYISTSVH 180
QY      181 HVLWIMHCFVTVLPSCIFPILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
        181 HVLWIMHCFVTVLPSCIFPILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
DB      181 HVLWIMHCFVTVLPSCIFPILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240
QY      241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINPFLYCFISGRFTMAAATL 300
        241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINPFLYCFISGRFTMAAATL 300
DB      241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINPFLYCFISGRFTMAAATL 300
QY      301 KAFKCKQKQPVQPYTHNFSITSSPWPISPANSHCICMLVYQYDKNGKPIKVP 353
        301 KAFKCKQKQPVQPYTHNFSITSSPWPISPANSHCICMLVYQYDKNGKPIKVP 353
DB      301 KAFKCKQKQPVQPYTHNFSITSSPWPISPANSHCICMLVYQYDKNGKPIKVP 353

RESULT 10
US-10-689-832-22
; Sequence 22, Application US/10689832
; Publication No. US20040121380A1
; GENERAL INFORMATION:
; APPLICANT: Majumder, Kamud
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729DIV1
; CURRENT APPLICATION NUMBER: US/10/689,832
; CURRENT FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-689-832-22
```

Query Match 99.8%; Score 1868; DB 16; Length 353;
Best Local Similarity 99.7%; Pred. No. 5,3e-163;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MEHTHAHLAANSSLSWMSQSAAGFVAVVYYSLLCIGLPANILTVIISQVARROK 60
Db 1 MEHTHAHLAANSSLSWMSQSAAGFVAVVYYSLLCIGLPANILTVIISQVARROK 60
QY 61 SSYNVLLAALADIVLFFIVFVDFLEDFILNMOPQVDPKIIIEVLEFSSIHSTWTV 120
Db 61 SSYNVLLAALADIVLFFIVFVDFLEDFILNMOPQVDPKIIIEVLEFSSIHSTWTV 120
QY 121 PLTIDRYIAVCHPLKHTVSPARTKRVIVSVYITCFILTSIPYWMPIWTEEDYISTSVH 180
Db 121 PLTIDRYIAVCHPLKHTVSPARTKRVIVSVYITCFILTSIPYWMPIWTEEDYISTSVH 180
QY 181 HVLIMHCTVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISFATL 240
Db 181 HVLIMHCTVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISFATL 240
QY 241 WAPRIMILYHLYGAPIONRWLVHMSDIANMLALNTAINFPLXCFISKRPRTMAATL 300
Db 241 WAPRIMILYHLYGAPIONRWLVHMSDIANMLALNTAINFPLXCFISKRPRTMAATL 300
QY 301 KAFKFCQKQPVQFYTNHNSITSSPWISPAISHCIRKMLVYQYDNGKRPDK 353
Db 301 KAFKFCQKQPVQFYTNHNSITSSPWISPAISHCIRKMLVYQYDNGKRPDK 353
```

RESULT 11

US-09-813-432-60

/ Sequence 60, Application US/09813432
/ Publication No. US20030148485A1
/ GENERAL INFORMATION:
/ APPLICANT: Taupier Jr., Raymond J
/ APPLICANT: Majmuder, Kamud
/ APPLICANT: Spaderna, Steven K
/ APPLICANT: Smithson, Glenda
/ APPLICANT: Mezes, Peter S
/ APPLICANT: Verne, Corine A. M.
/ TITLE OF INVENTION: No. US20030148485A1 Polypeptides and Amino Acids Encoding Same
/ FILE REFERENCE: 15966-729
/ CURRENT APPLICATION NUMBER: US/09/813,432
/ PRIOR FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,836
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/193,843
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 60
/ LENGTH: 350
/ TYPE: PRT
/ ORGANISM: Homo sapiens

US-09-813-432-60

Query Match 99.2%; Score 1857; DB 10; Length 350;
Best Local Similarity 100.0%; Pred. No. 5,4e-162;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MEHTHAHLAANSSLSWMSQSAAGFVAVVYYSLLCIGLPANILTVIISQVARROK 60
Db 1 MEHTHAHLAANSSLSWMSQSAAGFVAVVYYSLLCIGLPANILTVIISQVARROK 60
QY 61 SSYNVLLAALADIVLFFIVFVDFLEDFILNMOPQVDPKIIIEVLEFSSIHSTWTV 120
Db 61 SSYNVLLAALADIVLFFIVFVDFLEDFILNMOPQVDPKIIIEVLEFSSIHSTWTV 120
QY 121 PLTIDRYIAVCHPLKHTVSPARTKRVIVSVYITCFILTSIPYWMPIWTEEDYISTSVH 180
Db 121 PLTIDRYIAVCHPLKHTVSPARTKRVIVSVYITCFILTSIPYWMPIWTEEDYISTSVH 180
QY 181 HVLIMHCTVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISFATL 240
Db 181 HVLIMHCTVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISFATL 240
QY 241 WAPRIMILYHLYGAPIONRWLVHMSDIANMLALNTAINFPLXCFISKRPRTMAATL 300
Db 241 WAPRIMILYHLYGAPIONRWLVHMSDIANMLALNTAINFPLXCFISKRPRTMAATL 300
QY 301 KAFKFCQKQPVQFYTNHNSITSSPWISPAISHCIRKMLVYQYDNGKRPDK 350
Db 301 KAFKFCQKQPVQFYTNHNSITSSPWISPAISHCIRKMLVYQYDNGKRPDK 350
```

RESULT 12

US-10-174-364-60

/ Sequence 60, Application US/10174364
/ Publication No. US20030216308A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson et al.
/ TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-729CIP2
/ CURRENT APPLICATION NUMBER: US/10/174,364
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,836
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/193,843
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 128
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 60
/ LENGTH: 350
/ TYPE: PRT
/ ORGANISM: Homo sapiens

US-10-174-364-60

Query Match 99.2%; Score 1857; DB 14; Length 350;
Best Local Similarity 100.0%; Pred. No. 5,4e-162;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY. 121 PLTIDRYIACHPKHYTVSPAR

```
Db 121 PLTIDRYIAVCHPLKXHTVSPARTKRVISVYITCFELTSPYWMNWTEDYISTSVH 180
QY 181 HVLIMHCFVTVYVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
Db 181 HVLIMHCFVTVYVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
QY 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFELYCFISKREPTMAAATL 300
Db 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFELYCFISKREPTMAAATL 300
QY 301 KAFKCOQOPVOFYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIK 350
Db 301 KAFKCOQOPVOFYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIK 350

RESULT 15
US-09-813-432-20
/ Sequence 20, Application US/09813432
/ Publication No. US20030148485A1
/ GENERAL INFORMATION:
/ APPLICANT: Taupier Jr., Raymond J
/ APPLICANT: Majmuder, Kamud
/ APPLICANT: Spaderna, Steven K
/ APPLICANT: Smithson, Glenda
/ APPLICANT: Mezes, Peter S
/ APPLICANT: Vernet, Corine A. M.
/ TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
/ FILE REFERENCE: 15966-729
/ CURRENT APPLICATION NUMBER: US/09/813,432
/ CURRENT FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,836
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/193,843
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 20
/ LENGTH: 372
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-813-432-20

Query Match 99.2%; Score 1857; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 5.8e-162;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHAAANSLSWMSGACGLGFVPPVYVYSLILCLGLPANTLITVYIISQLVARROK 60
Db 1 MEHTAHAAANSLSWMSGACGLGFVPPVYVYSLILCLGLPANTLITVYIISQLVARROK 60
QY 61 SSYNTLALAAADIIVLFVIVDFLLEDFILNMQMPQVDPDKIEVLEPSSIHSTISWITV 120
Db 61 SSYNTLALAAADIIVLFVIVDFLLEDFILNMQMPQVDPDKIEVLEPSSIHSTISWITV 120
```

```
QY 121 PLTIDRYIAVCHPLKXHTVSPARTKRVISVYITCFELTSPYWMNWTEDYISTSVH 180
Db 121 PLTIDRYIAVCHPLKXHTVSPARTKRVISVYITCFELTSPYWMNWTEDYISTSVH 180
QY 181 HVLIMHCFVTVYVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
Db 181 HVLIMHCFVTVYVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
QY 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFELYCFISKREPTMAAATL 300
Db 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFELYCFISKREPTMAAATL 300
QY 301 KAFKCOQOPVOFYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIK 350
Db 301 KAFKCOQOPVOFYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIK 350
```

Search completed: December 14, 2004, 20:27:31
Job time : 50.3994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2004, 20:27:39, Search time 3552 Seconds

(without alignments)
4699.680 Million cell updates/sec

Title: US-09-995-225B-16

Perfect score: 1872
Sequence: 1 MHTAHALANSLSNWSPG.....CIKMLVYDYDKGKPIKVP 353

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q/cgcn2.1/USPTO.spool/US0999225/rnat.14122004.110733.18024/app.query.fasta.1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=gcg -MIMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0999225.QCGN.1.1.2527.rnat.14122004.110733.18024 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.1.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sbs.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1872	100.0	1059	6	AX463227 Sequence
2	1872	100.0	1062	6	BD183203 Novel G p
3	1872	100.0	1062	6	BD183204 Novel G p
4	1872	100.0	1062	6	AX498192 Sequence

Result No.	Score	Query Match	Length	DB ID	Description
5	1872	100.0	1062	6	BD103851
6	1872	100.0	1062	6	BD103852
7	1872	100.0	1167	6	AY635179
8	1868	99.8	1343	6	AX254977
9	1857	99.2	1119	6	AX254975
10	1857	99.2	1130	6	AX375235
11	1844	98.5	1062	6	AX664703
12	1844	98.5	1062	6	AX664701
13	1747	93.3	1062	6	AX463235
14	1641.5	87.8	957	6	CQ739127
15	1641.5	87.7	1788	6	AX375442
16	1638	87.5	1002	6	AX463231
17	1638	87.5	1032	6	AX463230
18	1638	87.5	1070	6	BD144286
19	1638	87.5	1158	6	AB083594
20	1638	87.5	1167	9	AB083594
21	1638	87.5	1826	6	AX463232
22	1638	87.5	101882	6	AC021089
23	1638	87.5	110000	2	AL139235_0
24	1631	87.1	1000	6	AX342465
25	1599.5	85.4	18336	10	AC124438
26	1598.5	85.4	140885	2	AC119798
27	1598.5	85.4	203372	2	AC121346
28	1598.5	85.4	230813	2	AC114097
29	1471	78.6	864	6	AX330119
30	1421	75.9	963	6	AX657444
31	1317	70.4	795	6	AX549301
32	1317	70.4	795	6	HS403165
33	1277.5	68.2	1149	5	AY288414
34	994	53.1	558	9	AY255545
35	912	48.7	930	6	AX230116
36	773	41.3	1113	5	AY288413
37	737	39.4	945	9	AY255622
38	737	39.4	1122	6	BD185294
39	737	39.4	1122	6	BD140885
40	737	39.4	1125	6	AX498182
41	737	39.4	1189	9	AY288421
42	737	39.4	1962	6	AX375239
43	733	39.2	1125	6	AX644706
44	733	39.2	1125	6	AX664704
45	730.5	39.0	951	6	AX657458

ALIGNMENTS

RESULT 1
AX463227
LOCUS
DESCRIPTION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

AX463227
Sequence 1 from Patent WO0248358.
AX463227.1 GI:21886199
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1
Smolyar, A., Zhu, Z., Encinae, J., Watanabe, S. and Okiyama, H.
Regulation of human chemokine-like receptor
Patent: WO 0248358-A 1 20-JUN-2002;
Bayer Aktiengesellschaft (DE)
Location/Qualifiers
1..1059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 3,94e-149
Score: 1872.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Length: 1059
Matches: 353
Conservative: 0
Mismatch: 0

Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-995-225b-16 (1-353) x AX463227 (1-1059)

```

QY 1 MetGluHisThrHisAlaHisIleuValAlaAsnSerIleuSerTrpTrpSerProGly 20
DB 1 ATGGAGCAGACGACGCCACCTCGCAGCCAAACAGCTGCTGCTGAGTGTCCCGGC 60
QY 21 SerIleAcysGlyLeuGlyPheValProValValIleTrpSerIleuLeuLeuCyseuGly 40
DB 61 TCGGCTGGCGGCTTGGGTTTGTGTCGCCGTGCTACTACAGCCTTGGCTGTGCTGCTGCT 120
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
DB 121 TTACACGAAATATCTTGAACAGATATATCTCTCCAGCTGGTGGCAGAAAGACAGAAAG 180
QY 61 SerSerTrpAsnTrpIleuLeuValAlaValAlaAspIleLeuValIlePhePheIle 80
DB 181 TCCTCTACAACTATCTCTTGGCAGCTGCTGCTGCGACATCTTGGCTCTTTTCATA 240
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 241 GTGTTTGGACATCTCTGTTGGAAGATTTCATCTTGAAACATGACAGATGCTCAGTCCCC 300
QY 101 AspLysIleIleGlnValIleuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
DB 301 GACAAAGATCATAGAGTGTGGAATCTCATCTCCACCTCCATATGATTTACTGTA 360
QY 121 ProLeuThrIleAspAspGlyIleAlaValCysHisProLeuLysTrpHisThrValSer 140
DB 361 CGGTAAACATTAACAGATATATCGTGTCTGACCCGCTCAAGTACACACAGGTCTCA 420
QY 141 TrpProAlaArgThrArgLysValIleValSerValIleHisCysPheLeuThrSer 160
DB 421 TACCCAGCCCGGACCCGGAAAGTCATTTGATTTTAACTACCTGCTCTCCGACACAC 480
QY 161 IleProTrpTrpTrpProAsnIleTrpThrGluAspTrpIleSerThrSerValHis 180
DB 481 ATCCCTATTAACAGTGTGGCCCAACATCTGAGTGAAGACTATACATGACACCTGTGAT 540
QY 181 HisValIleuIleTrpIleHisCysPheThrValIleLeuValProCysSerIlePhePhe 200
DB 541 CAGCTCTCATCTGATTCACATGCTGCTTACCGTCTACCTGCTGCTCCCTCTCTCTTC 600
QY 201 IleLeuAsnSerIleIleValIleTrpLysLeuArgArgLysSerAsnPheArgLeuArgIly 220
DB 601 ATCTTGAATCAATCATGTGTGTAACAAGCTCAGAGAGAAAGCAATTTTCGTCTCCGTGC 660
QY 221 TrpSerThrGlyLysThrThrAlaIleLeuPheThrIleHisSerIlePheAlaThrLeu 240
DB 661 TACTCCAGGGAGAACCCAGCCCATCTTGTTCACCATTAATCTTCATCTTGGCCACACT 720
QY 241 TrpAlaProArgIleIleMetIleLeuTrpHisLeuTrpGlyValAlaProIleGlnAsnArg 260
DB 721 TGGGCCCCCGCATATCATATGATTTCTTTCACCTCTTATGGGGCCGCCATCCAGAACGCG 780
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuValAlaLeuAsnThrAlaIle 280
DB 781 TGGCTGTGACATCATGTCCGACATTCGCAACATGCTAGCCCTTCTGAAACAGCCATC 840
QY 281 AsnPhePheLeuTrpCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300
DB 841 AACTTCTTCTCTACTGCTTCTATTCAGACAGGCTTCCGACCATGAGCGCCGACGCTC 900
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTrpHisAsnIleAsnPheSer 320
DB 901 AAGGTTTCTTCAAGTCCAGAGAACCTGTACAGTTCTTACACCAATCATTAATTTTCC 960
QY 321 IleHisSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTrp 340
DB 961 ATTAACAAGTACCCCTGGATCTCGCCGCAAACTCACATGACATCAAGATGCTGTGTAC 1020

```

QY 341 GlnTrpAspLysArgGlyLysProIleLysValSerPro 353
 DB 1021 CAGTATGACAAATAATGAAAACCTATTAAGATATCCCG 1059

RESULT 2
 BD183203 1062 bp DNA linear PAT 17-JUN-2003
 LOCUS Novel G protein-coupled receptor and its DNA.
 DEFINITION BD183203
 ACCESSION BD183203.1 GI:31875403
 VERSION JP 2002345481-A/3
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1062)

REFERENCE
 Terao, Y., Matsui, H. and Shintani, Y.
 Novel G protein-coupled receptor and its DNA
 Patent: JP 2002345481-A 3 03-DEC-2002;
 JOURNAL TAKEDA CHEMICAL INDUSTRIES LTD
 OS Homo sapiens (human)
 PN JP 2002345481-A/3
 PD 03-DEC-2002
 PF 01-JUN-2001 JP 2001166688
 PI YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
 PC C12N15/09, A61K45/00, A61P1/00, A61P3/00, A61P9/00, A61P25/00, PC
 A61P29/00,
 PC A61P35/00, A61P37/00, C07K14/705, C07K16/28, C12N1/15, C12N1/19, PC
 C12N1/21,
 PC C12N5/10, C12P21/02, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, PC
 G01N33/53,
 PC G01N33/566, C12N15/00, C12N5/00
 CC Novel G protein-coupled receptor and its DNA
 FH key
 FT source
 1..1062
 /location/Qualifiers
 /organism='Homo sapiens (human)'.
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 3.95e-149 Length: 1062
 Score: 1872.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-995-225b-16 (1-353) x BD183203 (1-1062)

```

QY 1 MetGluHisThrHisAlaHisIleuValAlaAsnSerIleuSerTrpTrpSerProGly 20
DB 1 ATGGAGCAGACGACGCCACCTCGCAGCCAAACAGCTGCTGCTGAGTGTCCCGGC 60
QY 21 SerIleAcysGlyLeuGlyPheValProValValIleTrpSerIleuLeuLeuCyseuGly 40
DB 61 TCGGCTGGCGGCTTGGGTTTGTGTCGCCGTGCTACTACAGCCTTGGCTGTGCTGCTGCT 120
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
DB 121 TTACACGAAATATCTTGAACAGATATATCTCTCCAGCTGGTGGCAGAAAGACAGAAAG 180
QY 61 SerSerTrpAsnTrpIleuLeuValAlaValAlaAspIleLeuValIlePhePheIle 80
DB 181 TCCTCTACAACTATCTCTTGGCAGCTGCTGCTGCGACATCTGTGCTCTTTTCATA 240
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 241 GTGTTTGGACATCTCTGTTGGAAGATTTCATCTTGAAACATGACAGATGCTTCCGCC 300

```

QY 101 AsgLyvlellellegluValleugluPheSerSerllelthSerSerllelthVal 120
 DB 301 GAGAAAGTCAATGAGTCTGAGATTCATCCATCCACACCTCCATATGATTAAGTCA 360
 QY 121 ProLeuThr11eAaPrgTyr11eAlaValCySH:SPProLeuTyr11eThrValSer 140
 DB 361 CGGTTAAACCATGACAGGTATATCGCGTCTGACACCGCTCAAGTACCAACAGGTCTCA 420
 QY 141 TyrProAlaArgThrArglyVal11eValSerValTyr11eThrCySPheLeuThrSer 160
 DB 421 TACCCAGCCCGCACCCGGAAAGTCATGTAAAGTTTACATCACTGCTTCGACACGAC 480
 QY 161 lleProTyrTyrTrpTrpProAenlleTrpThrGluAspTyr11eSerThrSerValHis 180
 DB 481 ATCCCTATTAATCGTGGCCCAACATCTGAGCTGAAGACTACATCAACACCTCTGTGCAT 540
 QY 181 HisValleu11eTrp11eHisCySPheThrVal11eValProCySerSer11ePhePhe 200
 DB 541 CAGCTCTATCTGGATCCACCTGCTTCAACGCTTACCTGCTGCTGCTCATCTTCTTC 600
 QY 201 11eLeuAaSer11e11eVal11eTyrLyLeuArgArgLySerAaPheArgLeuArgLy 220
 DB 601 ATCTTAACCTCAATCATTTGTGTACAAAGCTCAGAGGAAAGCAATTTTCGCTCCGTGAC 660
 QY 221 TyrSerThrGlyLySerThrAla11eLeuPheThr11eThrSer11ePheAlaThrLeu 240
 DB 661 TACTCCAGGGGAAAGCAACCGGCACTTGTTCACCATTAACCTTCATCTTGGCACACTT 720
 QY 241 TrpAlaProArg11e11eMet11eLeuTyr11eLeuTyr11eValPro11eGlnAaArg 260
 DB 721 TGGGCCCCCGGCACTCATCATGATTTCTTTACCACTCTATGGGGCGCCCAACGAAACGC 780
 QY 261 TrpLeuValHis11eMetSerAaP11eAlaAaMetLeuAlaLeuAaThrAla11e 280
 DB 781 TGGCTGTACACATCATGTCGCAATTCGCAACATGCTACCCCTTGGAAACAGACATC 840
 QY 281 AaPhePheLeuTyrCySPhe11eSerLyAaArgPheArgThrMetAla11eAlaThrLeu 300
 DB 841 AACTCTCTCTCACTGCTTCATCAGCAAGCGGTTCGCAACATGGAAGCGGCACGACCTC 900
 QY 301 LysAlaPhePheLySCySglnLySglnProValGlnPheTyrThrAaHisAaPheSer 320
 DB 901 AAGGCTTTCTTCAAGTGCAGAAAGCAACCTGTACAGTTCACACATCATTACTTTTCC 960
 QY 321 11eThrSerSerProTrp11eSerProAlaAaSerHisCyS11eLyMetLeuValTyr 340
 DB 961 ATTAACAAGTACGCCCTGATCTCGCGGCAAACTCACATGATCAAGAGTGTGTAC 1020
 QY 341 GlnTyrAspLyAaSnGlyLysPro11eLyValSerPro 353
 DB 1021 CAGTATGACAAATAAGAAAACCTATTAAGATATCCCG 1059

RESULT 3
 BD183204 1062 bp DNA 1linear PART 17-JUN-2003
 LOCUS Novel G protein-coupled receptor and its DNA.
 DEFINITION BD183204
 ACCESSION BD183204.1 GI:31875404
 VERSION JP 2002345481-A/4.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1062)
 Terao, Y., Matsui, H. and Shintani, Y.
 Novel G protein-coupled receptor and its DNA
 Patent: JP 2002345481-A 4 03-DEC-2002;
 TAKEDA CHEMICAL INDUSTRIES LTD
 OS Homo sapiens (human)
 PN JP 2002345481-A/4
 PD 03-DEC-2002
 PF 01-JUN-2001 JP 2001166688
 PI YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI

PC C12N15/09, A61K45/00, A61P1/00, A61P3/00, A61P9/00, A61P25/00, PC
 A61P29/00, PC
 PC A61P35/00, A61P37/00, C07K14/705, C07K16/28, C12N1/15, C12N1/19, PC
 C12N1/21,
 PC C12N5/10, C12P21/02, C12Q1/58, G01N33/15, G01N33/50, G01N33/53, PC
 G01N33/53,
 PC G01N33/566, C12N15/00, C12N5/00
 CC Novel G protein-coupled receptor and its DNA
 FH Key Location/Qualifiers
 FT Source 1. 1062
 FT Location/Qualifiers
 source 1. 1062
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 3,956-149 Length: 1062
 Score: 1872.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-995-225B-16 (1-353) x BD183204 (1-1062)

QY 1 MetGluHisThrHisAlaHisLeuAlaAaSerSerLeuSerTrpTrpSerProGly 20
 DB 1 ATGAGACACACGACGCCCACTCGGACGCAACAGTCTGCTGTGGTGTCCCGCCGCG 60
 QY 21 SerAlaCyGlyLeuGlyPheValProValTyrTyrSerLeuLeuCyLeuGly 40
 DB 61 TCGGCTGGGCGCTTGGGTTTCGGCCCGGTGCTTACTACAGCTCTTGGCTGTGCTCGGT 120
 QY 41 LeuProAlaAaThr11eLeuThrVal11e11eLeuSerGlyLeuValAlaArgArgGlnLys 60
 DB 121 TTACCAAGAAATATCTTGACAGATCATCTCTCCAGTGGTGGCAAGAAACAGAG 180
 QY 61 SerSerTyrAaThrLeuLeuAlaLeuAlaAaAsp11eLeuValLeuPhePhe11e 80
 DB 181 TCCTCTTACACATATCTCTTGGACCTCGCTGCGCGGACATCTTGGTCTCTTTTCATA 240
 QY 81 ValPheValAaPheLeuLeuGluAaPhe11eLeuAaMetGlnMetProGlnValPro 100
 DB 241 GTGTGTGAGACTTCTGTTGGAAGATTCATCTTGAACATGACAGATGCTCAGGTCCCC 300
 QY 101 AsgLyvlellellegluValleugluPheSerSerllelthSerSerllelthVal 120
 DB 301 GACAAAGTCAATGAAAGTCTGGAATTCATCATCCATCCACACCTCCATATGATTAAGTCA 360
 QY 121 ProLeuThr11eAaPrgTyr11eAlaValCySH:SPProLeuTyr11eThrValSer 140
 DB 361 CGGTTAAACCATGACAGGTATATCGCTCTGACACCGCTCAAGTACCAACAGGTCTCA 420
 QY 141 TyrProAlaArgThrArglyVal11eValSerValTyr11eThrCySPheLeuThrSer 160
 DB 421 TACCCAGCCCGCACCCGGAAAGTCATGTAAAGTTTACATCACTGCTTCGACACGAC 480
 QY 161 lleProTyrTyrTrpTrpProAenlleTrpThrGluAspTyr11eSerThrSerValHis 180
 DB 481 ATCCCTATTAATCGTGGCCCAACATCTGAGCTGAAGACTACATCAACACCTCTGTGCAT 540
 QY 181 HisValleu11eTrp11eHisCySPheThrVal11eValProCySerSer11ePhePhe 200
 DB 541 CAGCTCTATCTGGATCCACCTGCTTCAACGCTTACCTGCTGCTGCTCATCTTCTTC 600
 QY 201 11eLeuAaSer11e11eVal11eTyrLyLeuArgArgLySerAaPheArgLeuArgLy 220
 DB 601 ATCTTAACCTCAATCATTTGTGTACAAAGCTCAGAGGAAAGCAATTTTCGCTCCGTGAC 660
 QY 221 TyrSerThrGlyLySerThrAla11eLeuPheThr11eThrSer11ePheAlaThrLeu 240

```

Db      661 TACTCAGCGGGAAGACACCGCCATCTTGTACCATTACTTCATCTTGGCCACACTT
QY      241 TTPALPFAAGTlletlleuYrHlsleuYrGlyAlaProIlleGlnAsnArg
Db      721 TGGGCCCCCGCATCATCATGATGATCTTAAACACTCTAATGGGGCCCATCCAGACCGC
QY      261 TTPLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle
Db      781 TGGCTGTGTGCATCATCATGTCGACATTCGACACATGCTAGCCCTTGTGACACAGCCATC
QY      281 AsnPhelLeuYrCysPheIleSerIleSerIlePheArgPheArgThrMetAlaAlaThrLeu
Db      841 AACTCTCTCTCTACTACTGCTTCATCAGCAAGCGGTTCCGACATGGCCAGCCGCTC
QY      301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheYrThrAsnHisAsnPheser
Db      901 AAGGCTTTCTTCAAGTGTCCAGAGCAACCTGTACAGTTCTACACCAATCATTAATTCTTCC
QY      321 lIethSerSerProTIPrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr
Db      961 ATAACAAGTACCCCTGGATCTCGCGGCAAACTCACTGTCATCAAGATGCTGGTATC
QY      341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
Db      1021 CAGTATGACAAAATGAAAACCTATAAAGTATCCCCG 1059

RESULT 4
AX498192      1062 bp      DNA      linear      PAT 26-SEP-2002
LOCUS      Sequence 15 from Patent WO242461.
DEFINITION      AX498192
ACCESSION      AX498192.1 GI:23343117
VERSION      AX498192.1 GI:23343117
KEYWORDS
SOURCE      synthetic construct
ORGANISM      synthetic construct
            artificial sequences.
REFERENCE      1
AUTHORS      Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.
TITLE      Endogenous and non-endogenous versions of human g protein-coupled
            receptors
            Patent: WO 0242461-A 15 30-MAY-2002;
            Arena Pharmaceuticals, Inc. (US)
JOURNAL      Location/Qualifiers
FEATURES
            1..1062
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Novel Sequence"
ORIGIN
Alignment Scores:
Pred. No.:      3,95e-149      Length:      1062
Score:      1872.00      Matches:      353
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
            Gaps:      0
US-09-995-225B-16 (1-353) x AX498192 (1-1062)
QY      1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTIPrIleSerProGly 20
Db      1 ATGGAGCACACGACGCGCCACCTGTGCACGACCAAGCTGTGCTTGGTGGTCCCGCGC 60
QY      21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCysLeuGly 40
Db      61 TGGCTGTGTGCATCATCATGTCGACATTCGACACATGCTAGCCCTTGTGACACAGCCATC
QY      1 LeuProAlaAsnIleLeuThrValIleIleuSerGlnLeuValAlaArgArgGlnLys 60
Db      121 TTACACGCAAAATATCTTACAGTATCATCTCTCCAGCTGGTGGCAAGAGCAGAG 180

```

```

QY      61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db      181 TCCTCTCAACAATATCTTGGCAGCTGTGGCGGACATCTGTGCTCTCTTTTTCATA 240
QY      81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      241 GTGTGTGGACTTCTGTGGAGATTTCACTTGAACATCCAGATGCTCAGGTCCC 300
QY      101 AspLysIleIleGlnValLeuGlnPheSerSerIleHisThrSerIleTIPrIleThrVal 120
Db      301 GACAAATCATAGATGTGGAAATTCATCATCCATCCACCTCCATATGATGATCTGTA 360
QY      121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysThrHisThrValSer 140
Db      361 CGGTAAACATGACAGGATATCCCTGTGTGCCACCGCTCATATACACAGGCTTCA 420
QY      141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db      421 TACCCAGCCCGCACCCGGAAGTCAATGTAGTGTATCATCACTGCTTCTTGACAGC 480
QY      161 lIleProYrTyrTIPrIleProAsnIleTIPrHcGluAspTyrIleSerThrSerValHis 180
Db      481 ATCCCTTATTAATGATGTGCGCCAAACATCTTGACTGAGACTAGACATACAGACCTGTGAT 540
QY      181 HisValLeuIleTIPrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db      541 CACGTCTCATCTGGATCACTGCTTCCGCTTACCGTACGCGTACGCGCTGTCTCATCTTCTC 600
QY      201 lIleLeuAsnSerIleIleValTyrLysLeuArgTyrLysSerAsnPhaGluArgGly 220
Db      601 ATCTTGAATCATCATTTGTGTACAGCTCAGGAGGAGCAATTTTCGTCCCGTGGC 660
QY      221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db      661 TACTCCAGCGGGAAGACACCGCATCTTGTACCATTAACCTTCATCTTGGCACACTT 720
QY      241 TTPALPFAAGTlletlleuYrHlsleuYrGlyAlaProIlleGlnAsnArg 260
Db      721 TGGGCCCCCGCATCATCATGATGATCTTAAACACTCTAATGGGGCCCATCCAGACCGC 780
QY      261 TTPLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
Db      781 TGGCTGTGTGCATCATCATGTCGACATTCGACACATGCTAGCCCTTGTGACACAGCCATC 840
QY      281 AsnPhelLeuYrCysPheIleSerIleSerIlePheArgPheArgThrMetAlaAlaThrLeu 300
Db      841 AACTCTCTCTCTACTACTGCTTCATCAGCAAGCGGTTCCGACATGGCCAGCCGCTC 900
QY      301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheYrThrAsnHisAsnPheser 320
Db      901 AAGGCTTTCTTCAAGTGTCCAGAGCAACCTGTACAGTTCTACACCAATCATTAATTCTTCC 960
QY      321 lIethSerSerProTIPrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db      961 ATAACAAGTACCCCTGGATCTCGCGGCAAACTCACTGTCATCAAGATGCTGGTATC 1020
QY      341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
Db      1021 CAGTATGACAAAATGAAAACCTATAAAGTATCCCCG 1059

RESULT 5
BD103851      1062 bp      DNA      linear      PAT 27-AUG-2002
LOCUS      Novel G protein-coupled receptor and its DNA.
DEFINITION      BD103851
ACCESSION      BD103851.1 GI:22649425
VERSION      WO 0194382-A/3.
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
            1 (bases 1 to 1062)
            Terao, Y., Matsuura, H. and Shintani, Y.

```

TITLE Novel G protein-coupled receptor and its DNA
JOURNAL Patent: WO 0194582-A 3 13-DEC-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI

COMMENT OS Homo sapiens (human)
PN WO 0194582-A/3
PD 13-DEC-2001
PR 01-JUN-2001 WO 2001JP004643
PR 02-JUN-2000 JP 00P 170446, 23-JUN-2000 JP 00P 194926 PI
YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/12, C12P21/02, C07K14/705, C07K16/28, A61K45/00, A61P25/00, A61P29/00,
PC A61P9/00, A61P35/00, A61P3/00, A61P37/00, A61P1/00, G01N33/566, PC G01N33/50,
PC G01N33/15
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
FT Location/Qualifiers
FT /organism='Homo sapiens (human)'.
/db_xref='taxon:9606'

ORIGIN

Alignment Scores:
Pred. No.: 3,95e-149 Length: 1062
Score: 1872.00 Matches: 353
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-995-225b-16 (1-353) x BD103851 (1-1062)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
DB 1 ATGGACACACGACGCCACCTCGACGACCAAGCTCGCTGTGGTGGTCCCGCGG 60
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
DB 61 TGGGCTGGGGCTGGGTTTCGTCGCCGCTGTACTACAGCTTCTGCTCCCTCGGT 120
QY 41 LeuProAlaAsnLeuLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGln 60
DB 121 TTAACGCAAAATATCTTGCACAGTATCATCTCCACAGTGGTGCAGAGAGAG 180
QY 61 SerSerTyrAsnTyrLeuAlaLeuAlaAlaAlaAspIleLeuValLeuPhePheIle 80
DB 181 TCCTCTACACATATCTTGGACCTCGCTGCCGACATCTTGGCTCTTTTCATA 240
QY 81 ValPheValAspPheLeuGlyLeuAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 241 GTGTTGTGGACTCTCTGTGGAGATTCATCTTGAACATGCATGCTCAGGTCCC 300
QY 101 AspLysIleIleGlyValLeuGlnPheSerSerIleHisThrSerIleTrpIleThrVal 120
DB 301 GACAAATCATAGAAATGTGCGAATCTCATCATCACAACCTCCATATGATTAATCTGTA 360
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
DB 361 CGGTACACATTCAGAGATATATCGCTGTGCACCGCTCAAGTACACACAGGTCTCA 420
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
DB 421 TACCCAGCCCGCACCCGGAAGTCACTTGAAGTTCATATCACTCCGCTTCGACACAG 480
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpHisGlyAspTyrIleSerThrSerValHis 180
DB 481 ATCCCTATTAATCGTGGGCCCAACATCTGAGTGAAGACATACTACAGCACTCTGAGAT 540
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200

DB 541 CAGTCTCATCTGGATTCACCTGCTTACCGCTGCTGCTGCTCATCTTCTTC 600
QY 201 IleLeuAsnSerIleIleValTyrTyrLeuArgArgLysSerAsnPheArgLeuArgGly 220
DB 601 ATCTTAATCATCATCTGTTGTAAGAGCTCAGAGAGAGAGCAATTTTGTCTCCGTGGC 660
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
DB 661 TACTCAGCGGAGAGCAACCGGCATCTGTTCACCATTAACCTCTTGTGCACACTT 720
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
DB 721 TGGGCCCCCGCATATCATGATTTCTTACCACTTATGGGGCGCCCATCCAGAACCCG 780
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAlaThrAlaIle 280
DB 781 TGGCTGTACACATCATGTCCGACATTGCCAATCATCTAGCCCTTGTGAACACAGCCATC 840
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
DB 841 AACTTCTTCTTACGCTTTCATCAGCAAGCCGTTCCGACCATGCGACCGCCGACGCTC 900
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 901 AAGCTTCTTCAAGTCCAGAGCAACTGTACACTTTCACCAATCATATCTTTTCC 960
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
DB 961 ATTAACAAGTAGGCCCTCGATCTCGCGGAACTCATGCACTGCAATGAATGCTGGTAC 1020
QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
DB 1021 CAGTATGACAAAATGAGAAACCTATPAAAGTATCCCG 1059

RESULT 6
BD103852 1062 bp DNA linear PAT 27-AUG-2002
LOCUS BD103852
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD103852
VERSION BD103852.1 GI:22649426
KEYWORDS MO 0194582-A/4.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Terao Y., Matsui H. and Shintani Y.
AUTHORS Novel G protein-coupled receptor and its DNA
TITLE Patent: WO 0194582-A 4 13-DEC-2001;
JOURNAL TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI

COMMENT OS Homo sapiens (human)
PN WO 0194582-A/4
PD 13-DEC-2001
PR 01-JUN-2001 WO 2001JP004643
PR 02-JUN-2000 JP 00P 170446, 23-JUN-2000 JP 00P 194926 PI
YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/12, C12P21/02, C07K14/705, C07K16/28, A61K45/00, A61P25/00, A61P29/00,
PC A61P9/00, A61P35/00, A61P3/00, A61P37/00, A61P1/00, G01N33/566, PC G01N33/50,
PC G01N33/15
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
FT Location/Qualifiers
FT /organism='Homo sapiens (human)'.
/db_xref='taxon:9606'

ORIGIN

Alignment Scores:

Pred. No.: 3,95e-149 Length: 1062
 Score: 1872.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-995-225b-16 (1-353) x BD103852 (1-1062)

```

QY      1 MetGluHsHrThrsAlaHisAlaAlaAnSerLeuSerTrpTrpSerProGly 20
DB      1 ATGAGACACACGACGCCACCTCGCAGCAACAGCTCGCTGTGTGTGTCGCCGCG 60
QY      21 SerAlaCyGylLeuGlyPheValProValTyrTyrSerLeuLeuCyLeuGly 40
DB      61 TCGGCGCTGGGCTTGGGTTTCGTGCGCGGTCTACTACAGCCCTTGTGCTCGGT 120
QY      41 LeuProAlaAniLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
DB      121 TTACACGACAAATATCTTGAACAGTATCATCTCTCCAGCTGTGGCAGAGACAGAG 180
QY      61 SerSerTyrAsnTyrLeuLeuAlaAlaAlaAspIleLeuValLeuPhePheIle 80
DB      181 TCTCTCTACAACTATCTTGGCAGCTCGCTGCCAGCATCTTGGCTCTTTTCATA 240
QY      81 ValPheValAspPheLeuGluAlaAspPheIleLeuAnSerGlnMetProGlnValPro 100
DB      241 GTTGTGTGACATCTCTGTGGAGAGATTTCATTTTAACTAGAGATGCTCAGGTCGCC 300
QY      101 AspLysIleIleGlnValLeuGlnPheSerSerIleHisThrSerIleTrpIleThrVal 120
DB      301 GACAAAGATCATAGAAAGTGTGGAATCTCATCATCCACACCCCATATGAGATTACTGTA 360
QY      121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
DB      361 CGGTAAACATTCAGCAGGTATGCTGTCTGCGACCGCTCAGATACACACGCGTCTCA 420
QY      141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
DB      421 TACCCAGCCCGCAGCCCGGAAAGTCATTGTAAGTTTATCATCACTGCTCTTCCAGCCAC 480
QY      161 IleProTyrTyrTrpTrpProAniLeuTrpGlnAspTyrIleSerThrSerValHis 180
DB      481 ATCCCTATTACTGTGTGGGCCCAACATCTGAGCTGAAGACTACATCAGCAGCTGTGAT 540
QY      181 HisValIleuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
DB      541 CACGCTCATCTGATCCAGTCCAGTCTTCACCGTCTACCTGGTCCCTGCTCCATCTTCTTC 600
QY      201 IleLeuAnSerIleIleValTyrLysLeuArgLysSerAnPheArgLeuGly 220
DB      601 ATCTGAATCAATCATGTGTGTAACAGTCAGAGAGAGAGAGATTTCGTCCTCGCTGGC 660
QY      221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
DB      661 TACTCCAGGGGAGAACACCGCCATCTTGTTCACCATATCTCCATCTTGGCCACACTT 720
QY      241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnLysArg 260
DB      721 TGGGCGCCCGCATATCATGATTTTAAACCTCTTAAGGGGCGCCCATCCAGAGACGCG 780
QY      261 TrpLeuValHisIleMetSerAspIleAlaAnMetLeuAlaLeuAnThrAlaIle 280
DB      781 TGGCTGTGACATCATGTCCGACATTCGCAACATGCTAGCCCTTGTGAACAGCCATC 840
QY      281 AsnPheLeuTyrCysPheIleSerLysArgPheArgPheMetAlaAlaIleThrLeu 300
DB      841 AACTTCTCTCTACTGTCTTATCAGCAAGCGGTTCGACCATGGCAGCGCCACGCTC 900
QY      301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrTrpAniHisAnPheSer 320
  
```

```

DB      901 AAGGCTTTCTTCAAGTGCAGAGAACCTGTACAGTTCTTACACCAATCATATCTTTCC 960
QY      321 IleHrSerSerProTrpIleSerProAlaAnSerHisCysIleLeuMetLeuValTyr 340
DB      961 ATAAACAGTAAGCCCTCGATCTCGCGGAAACTCAGACTGCATCAAGATGCTGTGTAAC 1020
QY      341 GlnTyrAspLysAnGlyLysProIleLysValSerPro 353
DB      1021 CAGTATGACAAAATGAAAACCTATATAAAGTATCCCG 1059
  
```

RESULT 7
 AY635179 1167 bp mRNA linear PRI 04-JUL-2004
 LOCUS
 DEFINITION Homo sapiens G protein-coupled receptor 139 (GPR139) mRNA, complete cds.
 ACCESSION AY635179
 VERSION AY635179.1 GI:49413551
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1167)

Bonner, T.I., Nagle, J.W. and Kauffman, D.

Complete coding sequence of GPR139

Unpublished

2 (bases 1 to 1167)

Bonner, T.I., Nagle, J.W. and Kauffman, D.

Direct Submission

Submitted (25-MAY-2004) Lab of Genetics, NIMH, Bldg 36, Rm 3D06, MSC4094, Bethesda, MD 20892-4094, USA

FEATURES

SOURCE

1..1167
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16p12.3"
 /cfeature_type="brain"
 1..1167
 /gene="GPR139"
 92..1153
 /gene="GPR139"
 /codon_start=1
 /product="G protein-coupled receptor 139"
 /protein_id="AA165818.1"
 /db_xref="GI:49413552"
 /translation="MERTHAHLANSLSLWSPGSAAGLGFVPUVYVLLSLGCLPAN
 ILTVILISQVARROKSSVYVLLAAADILVLEFVDFLEDFLNMOMQVDPK
 IIEVLERSIHGSIWITVPLTIDRYAVCHPLKYHNVSPARTRKYIVSYRTCPILS
 IPIYVMPNITEDYVIGSVYHNLIMHCFVYVYVPCSPFIILSIIVYKLRKSNFRL
 RGVSTGKTTILITTSIFATLMAPRIIMLNLVGAPOQNRLVHIMSDIAMLALL
 NKLINFLVCFISKRFRITMAATLKAFKCKQKQVQVFTYTHNPSITSSPWISPANSHC
 KMLVQYDNGKPIKIVSP"
 complement(1142..1167)

CDS

gene

ORIGIN

complement(1142..1167)

Alignment Scores:

Pred. No.: 4.4e-149 Length: 1167
 Score: 1872.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-995-225b-16 (1-353) x AY635179 (1-1167)

```

QY      1 MetGluHsHrThrsAlaHisAlaAlaAnSerLeuSerTrpTrpSerProGly 20
DB      92 ATGAGACACACGACGCCACCTCGCAGCAACAGCTCGCTGTGTGTGTCGCCGCG 151
QY      21 SerAlaCyGylLeuGlyPheValProValTyrTyrSerLeuLeuCyLeuGly 40
  
```

Db 152 TGGCGTGGGGCTTGGGTTTCGTCGCCGTGTCTACTACAGCCTTTCGTGCTCGT 211
 QY 41 LeuProAlaAsnIleuThrValIleIleuSerGlnIleuValAlaArgArgGlnIly 60
 Db 212 TTACCGACAAATCTTGAAGATCATCTCTCCAGCTGGGAGAGACAGAG 271
 QY 61 SerSerTyrAsnTyrIleuLeuAlaAlaAlaAspIleuValIleuPhePheIle 80
 Db 272 TCTCTCAACTATCTCTGGCACTGCTGCGACATCTTGAGCTCTTTCAT 331
 QY 81 ValPheValAspPheLeuGluAspPheIleuAsnMetGlnMetProGlnValPro 100
 Db 332 GTGTTGTGGACTTCTCGTTGAAGATTTCATCTTGACATGACAGTCCAGCTCC 391
 QY 101 AspIleIleIleGluValIleuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 Db 392 GACAGATCATAGAGTGTGGAAATCTCATCTCATCCACCTTCATATGATTACTGTA 451
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuValTyrHisThrValSer 140
 Db 452 CCGTTAACATGACAGATATATGCTGTGTCACCGCTCAAGTACACAGGCTCTCA 511
 QY 141 TyrProAlaArgThrArgValIleValSerValTyrIleThrCysPheLeuThrSer 160
 Db 512 TACCGACCGCGACCCGGAAAGTCATTGTATAGTTTACATCACTGCTTCTGACACG 571
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
 Db 572 ATCCCTATATACGTGGCCCAACATCTGACATGAAACATCACTGACACTGTCAT 631
 QY 181 HisValIleuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 Db 632 CAGCTCTCATCTGAGATCCATGCTTCACTGCTACCTGCTGCTGCTCATCTTCTTC 691
 QY 201 IleLeuAsnSerIleIleValTyrTyrLeuArgArgLysSerAsnPheArgLeuArgLys 220
 Db 692 ATCTTAACATCATCTTGTATCAAGCTCAGAGAGAGCAATTTCTCTCCGTGCG 751
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleHisSerIlePheAlaThrLeu 240
 Db 752 TACTCCAGCGGAGAGACACCGCATCTGTTTCACTTACATTCATCTTGTGCACACTT 811
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValAlaProIleGlnAsnArg 260
 Db 812 TGGGCCCCCGCATCATGATTTCTTACCATCTTATGGGCGCCCATCCAGAACCGC 871
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetIleuAlaLeuAsnThrAlaIle 280
 Db 872 TGGCTGTACATCATGTCCTCCCATTTGCCAATCATCTAGCCCTTGTGAACACAGCCATC 931
 QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
 Db 932 AACTTCTCTCTACTGCTTCATCAGCAAGCGGTTCCGACACATGACAGCCGACAGCTC 991
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 Db 992 AAGGCTTTCTTCAAGGCCAGAGCAACCTGTACAGTTTACACATCATTAATCTTTCC 1051
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLeuMetLeuValTyr 340
 Db 1052 ATPACAGATGAGCCCGATCTGCGCGGCAATCTCACATGATCAAGATGCTGTGTC 1111
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 Db 1112 CAGTATGACAAATGAGAAACCTATATAAGTATCCCG 1150
 RESULT 8
 AX254977 1343 bp DNA linear PAT 10-OCT-2001
 LOCUS AX254977
 DEFINITION Sequence 21 from Patent WO0170978.
 ACCESSION AX254977
 VERSION AX254977.1 GI:16074504
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Taupier, R. J., Majumder, K., Spaderna, S. K., Smithson, G., Mezes, P. S.
 and Vernet, C. A.
 TITLE Polypeptides and nucleic acids encoding same
 JOURNAL Patent: WO 0170978-A 21 27-SEP-2001;
 Curren Corporation (US)
 FEATURES
 source
 1..1343
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,13e-148 Length: 1343
 Score: 1868.00 Matches: 352
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 1
 Query Match: 99.79% Indels: 0
 DB: 6 Gaps: 0
 US-09-995-225B-16 (1-353) x AX254977 (1-1343)
 QY 1 MetGluHisThrHisAlaHisAlaAsnSerLeuSerTrpTrpSerProGly 20
 Db 2 ATGAGACACACCGACCGCCACCTCGCAGCCACAGCTGCTGTGGTGTCCCGCG 61
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
 Db 62 TGGCGCTGGGCTTGGGTTTCGTCGCCGTGCTACTACAGCTCTTGTGCTCGCT 121
 QY 41 LeuProAlaAsnIleuThrValIleIleuSerGlnLeuValAlaArgArgGlnLys 60
 Db 122 TTACCGACAAATCTTGAAGATCATCTCTCCAGCTGTGGCAAGAGACAGAG 181
 QY 61 SerSerTyrAsnTyrIleuLeuAlaAlaAspIleuValIleuPhePheIle 80
 Db 182 TCTCTCACTATATCTTGTGACATCGCTGCGCATCATCTTGGTCTTTCAT 241
 QY 81 ValPheValAspPheLeuGluAspPheIleuAsnMetGlnMetProGlnValPro 100
 Db 242 GTGTTGTGGACTTCTGTGGAAGATTTCATCTTGACATGACATGCTCAGGCTCC 301
 QY 101 AspIleIleIleGluValIleuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 Db 302 GACAGATCATAGAGTGTGGAAATCTCATCTCATCCACCTCCATATGATTTACTGTA 361
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuValTyrHisThrValSer 140
 Db 362 CCGTTAACATGACAGATATATCACTGCTCCACCGCTCAGATACACAGGCTCTCA 421
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 Db 422 TACCGACCGCGACCGGAAAGTCAITGTATGTTTATCATCACTGCTTCCAGCAC 481
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
 Db 482 ATCCCTATATCTGTGTGCGCAATCTGATGAAAGTCACTACAGCACTCTGTGCT 541
 QY 181 HisValIleuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 Db 542 CAGCTCTCATCTGTGATCCATGCTTCACTGCTACCTGCTGCTCCATCTTCTTC 601
 QY 201 IleLeuAsnSerIleIleValTyrTyrLysLeuArgArgLysSerAsnPheArgLeuArgLys 220
 Db 602 ATCTTGAACTCATATGTGTACAGCTCAGAGAGAGCAATTTCTGCTCGTGGC 661
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleHisSerIlePheAlaThrLeu 240

Db 662 TACTCCAGGGAGAGACCAACCGCATCTTGTTCACATTACCTCCATCTTGGCACAATT 721
 Qy 241 TTPAIAProAgiIleIleMetIleLeuYrHisLeuYrGlyAlaProIleGlnAsnArg 260
 Db 722 TGGGCCCCCGCATCATCATCATCTTTTACCACTCTATGGGGCGCCCATCAGAACCGC 781
 Qy 261 TTPLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 Db 782 TGGCTGGACACATCATCATCTCCGACATTCGCAACATGCTAGCCCTTCTGAACACAGCCATC 841
 Qy 281 AsnPheLeuYrCysPheIleSerIleSerIlePheArgThrMetAlaAlaIleThrLeu 300
 Db 842 AACTCTCTCTCTCTCTCTCTCTCTCATCAGCAAGCGGTTCCGACACANTGGACCGCCACGCTC 901
 Qy 301 LysAlaPhePheLeuYrCysGlnIleGlnProValGlnPheYrThrAsnHisAsnPheSer 320
 Db 902 AAGCTTTCTTCAAGTGCAGAACCACTGTACAGTTCTACACCATCATTAATCTTTTCC 961
 Qy 321 IleThrSerSerProTPIIleSerProAlaAsnSerHisCysIleIleMetLeuValIle 340
 Db 962 ATACAGAGTACGCCCTCGATCTCGCGCAAACTCCACATCTGCATCAGATGCTGTGTAC 1021
 Qy 341 GlnTyrAspIleYrAsnGlyIleYrProIleYrValSerPro 353
 Db 1022 CAGTATGACAAAATGGAAAACTATTAAGTATCCCG 1060

RESULT 9
 AX254975 1119 bp DNA linear PAT 10-OCT-2001
 LOCUS Sequence 19 from Patent WO0170978.
 DEFINITION AX254975
 ACCESSION AX254975.1 GI:16074503
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS
 Taupier, R.U., Majumder, K., Spaderna, S.K., Smithson, G., Mezes, P.S.
 and Vernet, C.A.
 Polypeptides and nucleic acids encoding same
 Patent: WO 0170978-A 19 27-SEP-2001;
 Curagen Corporation (US)

FEATURES
 source
 1.1119
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 7.79e-148
 Score: 1857.00 Length: 1119
 Percent Similarity: 100.00% Matches: 350
 Best Local Similarity: 100.00% Conservative: 0
 Query Match: 99.20% Mismatches: 0
 DB: 6 Indels: 0
 Gaps: 0

US-09-995-225B-16 (1-353) x AX254975 (1-1119)

Qy 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
 Db 1 ATGGAGCACACGACGCCCACTCCGACGCACACACTGCTGTGTGTGTCTCCCGCGC 60
 Qy 21 SerAlaCysGlyLeuGlyPheValProValValIleYrYrSerSerLeuLeuCysLeuGly 40
 Db 61 TCGGCTTCGGCTTGGGTTCGTGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 120
 Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
 Db 121 TTACCAAGAAATATCTTACAGATATCATCTCTCCACGCTGTCGCAAGAACAGAACAG 180
 Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80

Db 181 TCTCTTACACATATCTTGGCAGTCGCTGTGCGACATCTTGGTCTCTTTTTCATA 240
 Qy 81 ValPheValAspPheLeuLeuGlyAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 Db 241 GTGTGTGTGACCTTCTGTGTGAGATTTTCATCTTGAACATGACATGACATGCTCAGGTCCC 300
 Qy 101 AspIleIleIleGlyValLeuGlyupheserSerIleHisThrSerIleTrpIleThrVal 120
 Db 301 GACAGATCATGAAAGTGTGAAATTCATCATCATCATCACTCCATATGATTAATCTGA 360
 Qy 121 ProLeuThrIleAspArgGlyIleAlaValCysHisProLeuYrYrHisThrValSer 140
 Db 361 CGGTATACCATATGACAGATATCGCTGTGCGACCGCGCTCAAGTACACAGGCTTCA 420
 Qy 141 TyrProAlaArgThrArgIleValIleValSerValIleIleThrCysPheLeuThrSer 160
 Db 421 TACCCAGCCCGCACCCGGAAGTCAATGTATAGTGTATCATCACCTGCTTCTTCAACAGC 480
 Qy 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGlyAspTyrIleSerThrSerValHis 180
 Db 481 ATCCCTATTAATGCTGTGCGCCCAACATCTGACCTGAGCTGACATCAGACACTCTGTGCAT 540
 Qy 181 HisValIleuIleTrpIleHisCysPheThrValIleValProCysSerIlePhePhe 200
 Db 541 CAGCTTCATCTGATCTGACCTGCTTCAACCTGATCCTGATGCTGCTGCTGCTGCTGCTGCT 600
 Qy 201 IleLeuAsnSerIleIleValIleYrIleLeuArgIleYrSerAsnPheArgLeuArgIle 220
 Db 601 ATCTTGAACCTGATCATCTGTGTACAGCTCAGAGAGAAAGACATTTTGTCTCGTGGC 660
 Qy 221 TyrSerThrGlyIleYrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 Db 661 TACTCCAGGGGAGAACACCGCCATCTGTTCACCATCTTCATCTTGGCACAATT 720
 Qy 241 TTPAIAProAgiIleIleMetIleLeuYrHisLeuYrGlyAlaProIleGlnAsnArg 260
 Db 721 TGGGCCCCCGCATCATCATCATCTTTTACCACTCTATGGGGCGCCCATCAGAACCGC 780
 Qy 261 TTPLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 Db 781 TGGCTGGACACATCATCATCTCCGACATTCGCAACATGCTTACCCCTTCTGAACACAGCCATC 840
 Qy 281 AsnPheLeuYrCysPheIleSerIleSerIlePheArgThrMetAlaAlaIleThrLeu 300
 Db 841 AACTCTCTCTCTCTCTCTCTCTCTCATCAGCAAGCGGTTCCGACACATGCGCGCACGCTC 900
 Qy 301 LysAlaPhePheLeuYrCysGlnIleGlnProValGlnPheYrThrAsnHisAsnPheSer 320
 Db 901 AAGCTTTCTTCAAGTGCAGAACCACTGTACAGTTCTACACCATCATTAATCTTTTCC 960
 Qy 321 IleThrSerSerProTPIIleSerProAlaAsnSerHisCysIleIleMetLeuValIle 340
 Db 961 ATACAGAGTACGCCCTCGATCTCGCGCAAACTCCACATCTGCATCAAGATGCTGTGTAC 1020
 Qy 341 GlnTyrAspIleYrAsnGlyIleYrProIleYrValSerPro 350
 Db 1021 CAGTATGACAAAATGGAAAACTATTAATA 1050

RESULT 10
 AX375235 1130 bp DNA linear PAT 01-MAR-2002
 LOCUS Sequence 25 from Patent WO0210387.
 DEFINITION AX375235
 ACCESSION AX375235
 VERSION AX375235.1 GI:19169953
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS
 Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R.,
 Elliot, V.S., Ramkumar, J., Baughn, M.R., Kallick, D.A., Walla, N.K.,

Hafajia, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Policky, J.L.,
Kearney, L., Grail, R.C., Warren, B.A. and Ding, L.
G-Protein coupled receptors
Patent: WO 0210387-A 25 07-FEB-2002;
Incyte Genomics, Inc. (US)

FEATURES

source 1. 1130
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474767CB1"

ORIGIN

Alignment Scores:

Pred. No.: 7.88e-148 Length: 1130
Score: 1857.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.20% Indels: 0
DB: 6 Gaps: 0

US-09-995-225b-16 (1-353) x AX375235 (1-1130)

```

QY 1 MetGluHsrThrsAlaHisLeuValAlaAsnSerSerLeuSerTrpTrpSerProGly 20
DB 12 ATGAGACACACGACGCCACCTCGACGCCAACGCTCGCTTGTGGTCCCGCCG 71
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCysLeuGly 40
DB 72 TCGGCGCTGGGCTTGTGTCGTCGCTGCTACTACAGCCCTTGTGCTGCTGCT 131
QY 41 LeuProAlaAsnLeuLeuThrValLeileuSerGlnLeuValAlaArgGlnLys 60
DB 132 TTACGACGAAATATCTTGACAGTGATCATCTCCAGCTGTCGACAGAGACAGAG 191
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspLeuValLeuPhePheIle 80
DB 192 TCCTCCACAACTATCTCTTGCGACCTGCTGCTGCCACATCTTGGCTCTTTTCATA 251
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 252 GTGTGTGTGACCTTCTGTTGAAAGATTTCATCTTGAACATGACGATGCTCAGGTC 311
QY 101 AspValLeileuGluValLeuGluPheSerSerIleIleThrsSerIleTrpIleThrVal 120
DB 312 GACAAAGATCATAGAGGCTGGAATTCATCTCCACACCTCCATATGATTAATCTGTA 371
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
DB 372 CCGTTAACCATTTGACAGGATATGCTGCTGCTGCCACCCGCTCAAGTACACACGGTCTCA 431
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
DB 432 TACCCAGCCCGCACCCGGAAAGTCATTTGAAGTTTAACTCACTCTCTCTGACACAGC 491
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
DB 492 ATCCCCATTAATACGTGTGCCCAACATCTGACGTGAACATCAACACACCTCTGTGCAT 551
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
DB 552 CAGCTCTCATCTGAGTCCACTGCTTCAACCTGCTCACTGCTGCTCTCTCACTTTCTTC 611
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
DB 612 ATCTTGAATCAATCATTTGTGTACAGCTCAGAGAGAGAACTTTTCTGCTCCGAGGC 671
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
DB 672 TACTCCACGGGAGAGACACCGCATCTTGTTCACATTAACCTCACTTGTCCACACTT 731
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260

```

```

DB 732 TGGGCCCCCGCATCATGATTTCTTACCACTCTGATGGGGGCCCATCCAGAACCGC 791
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
DB 792 TGGCTGTACACATCATATGCTCCGACATTCGACATTCGACATTCGACATTCGACATTC 851
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaIleThrLeu 300
DB 852 AACCTCTCTCTTCACTGCTTCACTGACAGAGGGTTCGACCAATGGACGCCGCGCTC 911
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 912 AAGCTTTCTTCAAGTCCAGAGCAACCTGTACATTCATCAACCAATCATTAATCTTTCC 971
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
DB 972 ATAAACAAGTACCCCTGATCTCGCGGCAACTCACTGATCAAGATGCTGTGTAC 1031
QY 341 GlnTyrAspLysAsnGlyLysProIleLys 350
DB 1032 CAGTATGACAAATAATGAAAACCTATATAA 1061

```

RESULT 11

AX664703 1062 bp DNA linear PAT 22-MAR-2003
LOCUS
DEFINITION Sequence 12 from Patent WO02074960.
ACCESSION AX664703
VERSION AX664703.1 GI:29164463
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1 Leiby, K.R., Kapeller-Libermann, R. and Glucksmann, M.
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE
JOURNAL
38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
of human proteins and uses thereof
Patent: WO 02074960-A 12-26-SEP-2002;
Millennium Pharmaceuticals, Inc. (US)

FEATURES

source 1. 1062
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 9.22e-147 Length: 1062
Score: 1844.00 Matches: 347
Percent Similarity: 98.87% Conservative: 2
Best Local Similarity: 98.30% Mismatches: 4
Query Match: 98.50% Indels: 0
DB: 6 Gaps: 0

US-09-995-225b-16 (1-353) x AX664703 (1-1062)

```

QY 1 MetGluHsrThrsAlaHisLeuValAlaAsnSerSerLeuSerTrpTrpSerProGly 20
DB 1 ATGAGACACACGACGCCACCTCGACGCCAACGCTCGCTTGTGGTCCCGCCG 60
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCysLeuGly 40
DB 61 TCGGCGCTGGGCTTGTGTCGTCGCTGCTACTACAGCCCTTGTGCTGCTGCT 120
QY 41 LeuProAlaAsnLeuLeuThrValLeileuSerGlnLeuValAlaArgGlnLys 60
DB 61 TCGGCGCTGGGCTTGTGTCGTCGCTGCTACTACAGCCCTTGTGCTGCTGCT 120
QY 121 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 240
DB 181 TCCTCTTACAACTATCTTGTGACATCTGCTGCTGCCGACATCTTGTCTTTTCATA 240
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100

```


REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-09-995-225b-16 (1-353) x CQ737667 (1-957)

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 23601 06-SEP-2002,
PE Corporation (NY) (US)
Location/Qualifiers
1..957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

8.18e-130 Length: 957
1643.00 Matches: 312
99.36% Conservative: 0
99.36% Mismatches: 2
87.77% Indels: 0
Gaps: 0

QY 40 G1YLeuPProAlaAenIleLeuThrValIleIleuSerGlnLeuValAlaArgArgGln 59
Db 13 GGTTCGACGAATATCTTGACAGTGATCATCTCTCCAGCGGGGCAAGAACAG 72
QY 60 LysSerSerTyraAntyTyleuLeuAlaLeuAlaAlaAspIleLeuValLeuPhe 79
Db 73 AAGGCTCCCTCAACAATCTCTGCGACCTGCTGCGGACGACATCTTGCTTTC 132
QY 80 IleValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnVal 99
Db 133 ATAGTGTTCGACCTTCTGTTGAAAGATTATCTTGAACATGAGATGCTCGAGTC 192
QY 100 ProAspLysIleIleGluValLeuGluPheSerSerIleIleThrSerIleTyrIleThr 119
Db 193 CCGGACAGATCATAGAGTGTGGAATCTCATCATCAACATGAGATGCTCGAGTC 252
QY 120 ValProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrVal 139
Db 253 GTHCCGTTAACATTAACAGATATGCTGCTGTCACCGCTCAAGTACACAGGTC 312
QY 140 SerTyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThr 159
Db 313 TCATACCCAGCCCGGACCCGGAAAGTCATTTAGTGTTAATCACTGCTTCGAGCC 372
QY 160 SerIleProTyrTyrTyrProAsnIleTyrThrGluAspTyrIleSerThrSerVal 179
Db 373 AGCATCTCCCTATTACGTGGGCCCAACATCTGAGACTGAGACTACATCAGCACTCTG 432
QY 180 HisHisValLeuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhe 199
Db 433 CATCAAGTCTCATCTGGATCCAGTCTTCAACGCTCACTGCTGCTCCCTCCATCTTC 492
QY 200 PheIleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArg 219
Db 493 TTCATCTTGAATCATATCTGTGTGAAGCTCAAGAGGAAGAGCAATTTCTCTCCGT 552
QY 220 G1YTySerThrG1YLeuThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThr 239
Db 553 GGCATCTCAGGGGAAAGCAACCGCATCTGTTCACCATTAACCTCCATCTTGGCCACA 612
QY 240 LeuTPAlaProArgIleIleMetIleLeuTyrHisLeuTyrG1YAlaProIleG1YAsn 259
Db 613 CTTTGGGCCCCCGCATCATGATTTCTTTCACCTCTTAAGGGCCGCCATCCAGAAC 672
QY 260 ArgTPLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAla 279
Db 673 CGTGGCTGGTACATCATGTCCGACATGTCCAAACATGTAGCCCTTCTGAACAGGCC 732

QY 280 IleAsnPhenLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThr 299
Db 733 ATCAATCTTCTCTCTACTGCTTCTTCAACAGCGGTTCGACCATGCGACCCCGACG 792
QY 300 LeuLysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPh 319
Db 793 CTCGAAGCTTCTTCAAGGCGCAGAGCAACTGTACAGTCTTACACCAATCATACCTT 852
QY 320 SerIleThrSerSerProTyrIleSerProAlaAsnSerHisCysIleLysMetLeuVal 339
Db 853 TCCATTAACAAGTACGCCCTGAGTCTGCGCGCAAACTCACACTGCATCAAGATGCTGTG 912
QY 340 TyrGlnTyrAspLysAsnG1YLysProIleLysValSerPro 353
Db 913 TACAGTATGACAAAATGGAABAACTTAAAGATATCCCG 954

RESULT 15
CQ739127 1788 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION
Sequence 25061 from Patent WO02068579.
ACCESSION
CQ739127
VERSION
CQ739127.1 GI:42341363
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-09-995-225b-16 (1-353) x CQ739127 (1-1788)

Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 25061 06-SEP-2002,
PE Corporation (NY) (US)
Location/Qualifiers
1..1788
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

2.24e-129 Length: 1788
1641.50 Matches: 320
91.32% Conservative: 6
89.64% Mismatches: 10
87.69% Indels: 21
Gaps: 2

QY 9 AlaAlaAsnSerSerLeuSerTyrTyrSerProG1YSerAlaCysG1YLeuG1YPheVal 28
Db 742 TCAAGCTCAGCTTCTTCTCCACCCCGACCCCGACGG----- 777
QY 29 ProValValTyrTyrSerLeuLeuG1Y----- 40
Db 778 ---ATCTCTTCAACATGTTCTCGATGTTCCTGGGGGAAATCTACCTGTACAAAAGAAATG 834
QY 41 -----LeuProAlaAenIleLeuThrValIleIleuSerGlnLeuValAla 56
Db 835 GGAACCTCACTGTCGCGGGAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCA 894
QY 57 ArgArgGlnLysSerSerTyrAntyTyleuLeuAlaLeuAlaAlaAspIleLeuVal 76
Db 895 AGAAGACAGAAAGTCCCTTCAACAATATCTTGGCACTGCTGCTGCGCAATCTTGTC 954
QY 77 LeuPhePheIleValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMet 96
Db 955 CTTTTCATAGTGTTCGACCTTCTGTTGAAAGATTATCTTGAACATGCAAGATG 1014
QY 97 ProG1YValProAspLysIleIleGluValLeuGluPheSerSerIleIleThrSerIle 116

Db 1015 CCTCAGGTCCCGACAGATCATAGAGTGTGAAATTCTCATTCATCCACCTTCATA 1074
Qy 117 TTPILIEThValProleuThriIleAspArgTyrIleAlaValCysHisProleuLysTyr 136
Db 1075 TGGATTACTGTACCCGTTAACCAATTGACAGGTATATGCTGTGCGACCCGCTCAAGTAC 1134
Qy 137 HisThrValSerTyrProAlaArgThraGlyValIleValSerValTyrIleThrCys 156
Db 1135 CACAGGTCATACCCAGCCCGACCCGGAAGATCATGTAGAGTTTATCATCACCTGC 1194
Qy 157 PheLeuThrSerIleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSer 176
Db 1195 TTCCTGACAGCATCCCTTATCTGTGGCCCAACATCTGAGCTGAGACTACATCAGC 1254
Qy 177 ThrSerValHisHisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCys 196
Db 1255 ACCCTGTGCATCACCTCTCATCTGTGATCCACTGCTCACCGCTTACCTGTGCGCTGC 1314
Qy 197 SerIlePhePheIleleuAnserIleIleValTyrLysLeuArgArgLysSerAnphe 216
Db 1315 TCCATCTTCTCATCTTGAATCATATGATGTATACAGCTCAGAGAGAGCAATTTT 1374
Qy 217 ArgLeuArgGlyTyrSerThrGlyLysThrThraIleLeuPheThrIleThrSerIle 236
Db 1375 CGTCTCCGTGGCTACTCCACGGGGAAGACACCGCATCTGTTCACCATTACTTCATC 1434
Qy 237 PheAlaThrLeuTrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaPro 256
Db 1435 TTGGCCACACTTTGGGCCCCCGCATCATGATTTCTTTACCACTTATGGGGCGCCC 1494
Qy 257 IleGlnAsnArgTrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeu 276
Db 1495 ATCCAGAACCGCTGGCTGTACACATCATGTCCGACATTGCCAACATGCTAGCCCTTCTG 1554
Qy 277 AsnThrAlaIleAsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAla 296
Db 1555 AACACAGCATCACTTCTCTCTACTGCTTCATCAGCAAGCGGTTCCGCAACATGGCA 1614
Qy 297 AlaAlaThrLeuLysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsn 316
Db 1615 GCCGCCACGCTCAAGGCTTCTTCAAGTCCAGANAGCAACTGTACAGTTCTACACCAAT 1674
Qy 317 HisAsnPheSerIleThrSerSerProTyrPileserProAlaAsnSerHisCysIleLys 336
Db 1675 CATTACTTTTCAATAACAAGTAGCCCTGTGATCTGCCGCAAACTCACACTGCATCAAG 1734
Qy 337 MetLeuValTyrGlnTyrAspLysAangLysProIleLysValSerPro 353
Db 1735 ATGCTGTGTACCAAGTATGACAAAATGAAAACTTATAAAGTATCCCCG 1785

Search completed: December 14, 2004, 21:34:44
Job time : 3563 aeca

This Page Blank (uspto)

XX 14-DEC-2000; 2000US-0255150P.
 PR 02-APR-2001; 2001US-0280110P.
 PR 21-JUN-2001; 2001US-0299474P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Smolyar A, Zhu Z, Encinas J, Watanabe S, Okigami H;
 XX WPI; 2002-547858/58.
 DR P-P8DB; ABB79517.
 XX
 PT New isolated polynucleotide encoding a chemokine-like receptor
 PT polypeptide for treating e.g. asthma, myocardial infarction, human
 PT immunodeficiency virus infection, or chronic obstructive pulmonary
 PT disease.
 XX
 PS Claim 1; Fig 1; 114pp; English.

CC The present sequence is that of cDNA encoding a novel human chemokine-
 CC like receptor (see AMN79517) of 353 amino acids and having 7 putative
 CC transmembrane domains, consistent with the structure of a G-protein
 CC coupled receptor. Its closest human homologue is C-C chemokine receptor
 CC 3. The novel receptor is expressed at low levels in most tissues. It is
 CC expressed at a high level in phytohemagglutinin-stimulated CD8+ cells,
 CC but in none of the other immune cells tested. It may act as a receptor of
 CC chemottractant molecules on activated lymphocytes and be involved in
 CC cell trafficking and homing to sites of infection, inflammation or tissue
 CC injury. Regulation of activity of the novel receptor can therefore be
 CC used to treat cardiovascular, immunological and inflammatory diseases,
 CC including asthma and chronic obstructive pulmonary disease (COPD). The
 CC receptor may also be a target for viruses that reside in the nervous
 CC system. Regulating the binding of ligands, e.g. chemottractant molecules
 CC or virus particles, to the receptor can therefore be used to modulate the
 CC immune response to inhibit viral infections, including HIV infection. A
 CC claimed method of reducing activity of the receptor involves contacting a
 CC cell with a reagent (preferably an antibody, antisense oligonucleotide or
 CC ribozyme) to a product (preferably RNA or a polypeptide) encoded by a
 CC polynucleotide encoding the human chemokine-like receptor in vivo or in
 CC vitro. A claimed method of treating a chemokine-like receptor dysfunction
 CC related disease selected from HIV infection, a cardiovascular disorder,
 CC asthma or COPD uses a reagent that modulates a function of the receptor
 CC
 XX
 SQ Sequence 1059 BP; 240 A; 342 C; 206 G; 271 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,8e-179 Length: 1059
 Score: 1872.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-995-225B-16 (1-353) x ABN84269 (1-1059)

QY 1 MetGluHLeThrHiEAlAHISLeuAlaIaAsnSerSerLeuSerTrpTrpSerProGly 20
 DB 1 ATGGAGCAGCAGCAGCCCACTCCGACGACACACCTGCTGTGGTGGTCCCGCGC 60
 QY 21 SerAlaCyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
 DB 61 TCGGCTCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
 QY 41 LeuProAlaAsnLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
 DB 121 TTACCAAGAAATATCTTGAACAGTATCATCTCTCCACGCTGGTGGAGAAGACAGAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValIleLeuPhePhe 80
 DB 181 TCTCTCTACACTATCTCTTGGCACTCGCTCCGACATCTTGGTCTCTTTTATA 240
 QY 81 ValPheValaAspPheLeuLeuGlnaAspPheIleLeuAsnMetGlnMetProGlnValPro 100

DB 241 GTGTTGTGGACTTCCTGTTGGAATTTATCTTGAACATGACATGCTCAGTCCCC 300
 QY 101 AspLysIleIleGluValLeuGluPheSerSerIleIleThSerIleTrpIleThrVal 120
 DB 301 GACAGATCATAGAAATGCTGGAAATTCATCATCATCCACACCTCATATGATTACTGTA 360
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLeuThyHisThrValSer 140
 DB 361 CGTTAACATTTGACAGGATATCCGTGTGCCACCCGCTCAAGTACACAGGATCTCA 420
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGACCCGGAAGTCAATTAAGTTTACATCACTGCTTCGACCGC 480
 QY 161 IleProTyrTyrTrpTrpProAsnIleThrPheGluAspTyrIleSerThrSerValHis 180
 DB 481 ATCCCTTATCTAGTGGCCCAACATCTGACCTGAAGACTACATCAAGACACCTGTGAT 540
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 541 CACGCTTCATCTGATCACTGCTTCAACGCTTACCTGATCGCTGCTCATCTTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrIleLeuArgArgLysSerAsnPheArgLeuArgGly 220
 DB 601 ATCTTGAACATCATATGTGTACAGCTCAGGAGGAGACCAATTTTCGTCGCGGCG 660
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 661 TACTCCACGGGAGAGACACCGCATCTGTGTACCATTACTTCATCTTGGCACACTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 721 TGGGCCCCGACATCATATGATTTCTTACCATCTTATGGGCGCCATTCAGAACCGC 780
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
 DB 781 TGGCTGGTGAACATCATGTCGACATTCGCAACATGCTTACCTTCTGAAACAGCGCATC 840
 QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
 DB 841 AACCTCTTCCTCACTCCTTCATCAGCAAGCGGTCGCGACATGGACGCGCACAGCTC 960
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 901 AAGCTTTCTTCAAGTCCAGAACCAACCTGTACAGTCTACACCAATATTAATCTTTC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 DB 961 ATAAACAAGTAGCCCTGATCTCGCGGCAAACTCACTGATCAAGATGCTGCTGTAC 1020
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1021 CAGTATGACAAATAGAAAACCTATMAAAGTATCCCGC 1059

RESULT 2
 ABK15562
 ID ABK15562 standard, cDNA, 1062 BP.
 XX
 AC ABK15562;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE cDNA encoding novel G-protein coupled receptor TGR8, version #1.
 XX
 KW G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
 KW cytoskeletal; antiinflammatory; antitumor; fetal brain;
 KW central nervous system disease; circulatory organ disorder; cancer;
 KW metabolic disease; immunological disease; gastrointestinal disease;
 KW gene therapy; transgenic animal; human; gene, ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers


```
FT CDS 1..1062
FT /*tag= a
FT /product= "TGR8"
FT /note= "G-protein coupled receptor"
XX
XX MO200194582-A1.
XX
XX 13-DEC-2001.
XX
XX 01-JUN-2001; 2001MO-JP004643.
XX
XX 02-JUN-2000; 2000JP-00170446.
XX
XX 23-JUN-2000; 2000JP-00194926.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Terao Y, Matsui H, Shintani Y;
XX
XX MPI; 2002-164317/21.
XX
XX P-PSDB; AAU76416.
XX
XX Human fetal brain-originated G protein-coupled receptor protein TGR8 and
XX encoding DNA, for developing drugs to treat e.g. diseases of the central
XX nervous system or circulatory organs, cancer, and metabolic diseases.
XX
XX Claim 5; Page 94; 102pp; Japanese.
XX
XX The invention describes a human fetal brain-originated G protein-coupled
XX receptor protein, or its salt. The protein and encoded DNA are useful for
XX developing drugs to treat e.g. diseases of the central nervous system or
XX circulatory organs, cancer, metabolic diseases, immunological diseases
XX and gastrointestinal diseases. The invention also describes creation of a
XX probe for gene therapy and construction of a transgenic animal. This
XX sequence encodes the novel G-protein coupled receptor TGR8, described in
XX the method of the invention
XX
XX Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 6.83e-179 Length: 1062
XX Score: 1872.00 Matches: 353
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-09-995-225B-16 (1-353) x ABK15562 (1-1062)
XX
XX 1 MetGluHisThrHisAlaHisIleuAlaIleAsnSerLeuSerTyrProGly 20
XX |||||||
XX 1 ATGGAGCACACGACGCCACCTCGACGCAAGCTCGCTGGTGCTCCCGGC 60
XX |||||||
XX 21 SerAlaCysGlyLeuGlyPheValProValIleTyrTyrSerLeuLeuGly 40
XX |||||||
XX 61 TCGGCTGCGGCTTGGGTTTCGTCGCCGTGTACTACAGCTCTTGGCTCGCT 120
XX |||||||
XX 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGly 60
XX |||||||
XX 121 TTACCCAGCAAAATATCTTGACAGATCATCTCTCCAGCTGGTGCAAAAGAG 180
XX |||||||
XX 61 SerSerTyrAsnTyrIleuLeuAlaIleuAlaIleAspIleLeuValLeuPhe 80
XX |||||||
XX 181 TCTCTCTACAACTATCTTGGCACTGCTGCGCAGCAATCTTGGCTCTTTCA 240
XX |||||||
XX 81 ValPheValAspPheLeuGlyIleuAspPheIleLeuAsnMetGlnMetProGlnValPro 100
XX |||||||
XX 241 GGTGTTGTGGACTTCTCTGTTGGAAGATTTCATCTTGAACTGAGATGCTCAGGTCCC 300
XX |||||||
XX 101 AspIleIleIleGluValLeuGluPheSerSerIleHisThrSerIleTyrIleThrVal 120
XX |||||||
XX 301 GACAAATCATATAGAGTGTGAATTCATCATCCATCCACACCTCCATATGATTAAGTACTGTA 360
XX |||||||
XX 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLeuTyrHisThrValSer 140
```

```
DB 361 CCCTTAACCATTCACAGGATATCGCTGCTGCCACCGCTCAAGTACACAGCGTCTCA 420
XX |||||||
XX 141 TyrProAlaArgThrArgIleValIleValSerValTyrIleThrCysPheLeuThrSer 160
XX |||||||
XX 421 TACCCAGCCCGCACCCGAAAGCATTTGATGTTATCATCACTCGTCTCCGACGACG 480
XX |||||||
XX 161 IleProTyrTyrThrProAsnIleTyrPheGluAspTyrIleSerHisThrValHis 180
XX |||||||
XX 481 ATCCCTATTATGAGTGGCCCAACATCTGAGCTGAAGACTTACATCACTGACCTTGACAT 540
XX |||||||
XX 181 HisValIleuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
XX |||||||
XX 541 CAGCTCTCATCTCGATTCACCTCTTACCGGTACCGTCCGCTCCATCTTCTTC 600
XX |||||||
XX 201 IleLeuAsnSerIleIleValTyrIleValArgArgIleSerAsnPheArgLeuArgIle 220
XX |||||||
XX 601 ATCTTGAATCAATCATTTGTGACAGCTCAGAGGAGAAAGCAATTTGCTCTCCGTGGC 660
XX |||||||
XX 221 TyrSerThrGlyThrThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrIleu 240
XX |||||||
XX 661 TACTTCACGGGAGAGCACCGCATCTGTTCACCATTAACCTCCATCTTGGCACACTT 720
XX |||||||
XX 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
XX |||||||
XX 721 TGGGCCCCCGCATTCATCATGATTTCTTACCACTCTTATGGGGCGCCCATCCAGAACCGC 780
XX |||||||
XX 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuValAsnThrAlaIle 280
XX |||||||
XX 781 TGGCTGTGACATCATCATGTCGACATGTCACCATCTGAGCCCTTGAACACAGCATC 840
XX |||||||
XX 281 AsnPhePheLeuTyrCysPheIleSerIleValArgPheArgThrMetIleAlaIleThrLeu 300
XX |||||||
XX 841 AACTTTCTCTTACTGCTTCATCAGCAAGCGGTTCGCAACCTGGCAGCGCACGCTC 900
XX |||||||
XX 301 LysAlaPhePheLeuCysGlnIleValProValGlnPheTyrThrAsnHisAsnPheSer 320
XX |||||||
XX 901 AAGCTTTCTTCAAGTCCAGGAAAGCAACTGTACAGTTCTTACACCAATCACTTCTTCC 960
XX |||||||
XX 321 IleThrSerSerProTyrIleSerProAlaAsnSerHisCysIleIleMetLeuValTyr 340
XX |||||||
XX 961 ATTAACAAGTAGCCCGCTGATCTCGCGCAAACTCACATGATCAAGATGCTGCTGATC 1020
XX |||||||
XX 341 GlnTyrAspIleAsnGlyIleProIleValSerPro 353
XX |||||||
XX 1021 CAGTATGACAAAATGAAAAACCTATMAAGTATCCCG 1059
XX |||||||
XX
XX RESULT 3
XX ABK15563
XX ID ABK15563 standard; cDNA; 1062 BP.
XX
XX AC ABK15563;
XX
XX 08-MAY-2002 (first entry)
XX
XX cDNA encoding novel G-protein coupled receptor TGR8, version #2.
XX
XX G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
XX cytotoxic; antiinflammatory; antitumor; fetal brain;
XX central nervous system disease; circulatory organ disorder; cancer;
XX metabolic disease; immunological disease; gastrointestinal disease;
XX gene therapy; transgenic animal; human; gene; ss.
XX
XX Homo sapiens.
XX
XX OS
XX
XX Key Location/Qualifiers
XX CDS 1..1062
XX FT /*tag= a
XX FT /product= "TGR8"
XX FT /note= "G-protein coupled receptor"
XX
XX MO200194582-A1.
```


[illegible]

QY	221	TyrsSerThrgIyLysTrpTrpTrpAlaIleLeuPheThrTllePhrSerIlePheAlaTrpLeu	240
Db	661	TACTCCACGGGGAGACACCGCCATCTTTGTCACATTACCTTCATCTTGCCACACTT	720
QY	241	TrpAlaProArgrIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg	260
Db	721	TGGGCCCCCGGCATCATCATGATTTCTTTACCACTCTATGAGGGGCCCATCCAGAACGCG	780
QY	261	TrpLeuValHisIleMetSerArgIleAlaAsnMetLeuAlaLeuLeuAsnTrpAlaIle	280
Db	781	TGGCTGGTGACATCAATCATGCTCCGACATTGCCAACATGTCTACCCCTTCGAAACAGCGCATC	840
QY	281	AsnPhePheLeuTyrCysPheIleSerLeuArgPheArgThrMetAlaAlaIleThrLeu	300
Db	841	AACCTTCTTCTCTACCTCTCATCTTCATAGCAAGCGGTTCCGCAACCATGGACGCCCAAGCTC	900
QY	301	LysAlaPhePheLeuCySGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer	320
Db	901	AAGGCTTCTTCCAAAGTCCGACGAAGCAACGTACAGTTCTTACACCATCATTAACCTTTTC	960
QY	321	IleThrSerSerProTTPPIleSerProAlaAsnSerHisCysIleLeuMetLeuValTyr	340
Db	961	ATAACAAGATGAGCCCTCGATCTCGCCGCAAACTCACACATGCATCAAGATCGTGTGTAC	1020
QY	341	GlnTyrAspLysAsnGlyLysProIleLysValSerPro	353
Db	1021	CAGTATGACAAATAATGAGAAACCTATTATAAGTATCCCCG	1059
RESULT 5			
ID	ACC44115	standard, DNA; 1062 BP.	
AC	ACC44115;		
DT	27-JUN-2003	(first entry)	
DE	Human AXOR-57 coding sequence.		
XX	XX	Human AXOR-57 coding sequence.	
KM	ds; antibacterial; antifungal; antiviral; antiprotzoal; analgesic; gene;		
KM	cytostatic; antidiabetic; anorectic; anabolic; antisthmatic; antagonist;		
KM	antiparkinsonian; cardiac; hypotensive; hypertensive; nephrotoxic;		
KM	osteopathic; antiangiinal; antitumor; antiallergic; antiinflammatory;		
KM	antiemetic; nootropic; tranquilizer; neuroprotective; antidepressant;		
KM	vaccine; gene therapy; G-protein coupled receptor; receptor; infection;		
KM	pain; cancer; diabetes; obesity; anorexia; bulimia; asthma; stroke;		
KM	Parkinson's disease; osteoporosis; angina pectoris; ulcer; allergy;		
KM	vomiting; schizophrenia; depression; dementia; Huntington's disease.		
XX	OS	Homo sapiens.	
XX	XX		
FH	Key	Location/Qualifiers	
FT	CDS	1..1062	
FT	FT	/*tag= "B	
FT	FT	/product= "AXOR-57 protein"	
XX	XX		
PN	GB2365009-A.		
XX	XX		
PD	13-FEB-2002.		
PF	11-APR-2001; 2001GB-00009018.		
XX	XX		
PR	11-APR-2001; 2001GB-00009018.		
XX	XX		
PA	(SMIK) SMITHKLINE BEECHAM CORP.		
PA	(SMIK) SMITHKLINE BEECHAM PLC.		
XX	XX		
PI	Elahourbagy N, Gattu M, Michalovich D, Shabon U,		
XX	XX		
DR	WPI, 2003-203569/20.		
XX	XX		
DR	P-PSDB; ABP98724.		
XX	XX		
PT	New G-protein coupled receptor, AXOR 57, for diagnosing and treating		

PT diseases, such as, pain, cancer, diabetes, obesity, anorexia, asthma, Parkinson's disease, hypotension, hypertension, urinary retention, and osteoporosis.

XX Claim 2; Page 27, 32pp, English.

XX The invention relates to the isolation of the coding sequence of a novel CC G-protein coupled (7TM) receptor AXOR 57 (this sequence) or fragment CC encoding a polypeptide having at least 95% identity with AXOR-57 protein. CC The protein has been shown to have homology to the human chromosome 16 CC clone CIRM-H1_03A2 (AC008785). The protein and polynucleotides encoding CC the protein are used in screening for compounds that stimulate or inhibit CC the function or level of the polypeptide or polynucleotides, such as, CC agonists and antagonists. The protein, polynucleotides, and antibodies to CC the protein are used in diagnostic kits, to diagnose a disease. CC Polynucleotides encoding the protein are used for chromosome localization CC studies, or for tissue expression studies. The protein and nucleic acids CC encoding the protein are used in vaccines for treating diseases such as, CC bacterial, fungal, protozoal, and viral infections, pain, cancer, CC diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease, acute CC heart failure, hypotension, hypertension, urinary retention, CC osteoporosis, angina pectoris, myocardial infarction, stroke, ulcer, CC allergy, benign prostatic hypertrophy, migraine, vomiting, anxiety, CC schizophrenia, depression, delirium, dementia, severe mental retardation, CC or Huntington's disease

XX Sequence 1062 BP; 240 A; 342 C; 208 G; 272 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6 83e-179 Length: 1062
Score: 1872.00 Matches: 353
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-995-225b-16 (1-353) x ACC44115 (1-1062)

```

QY 1 MetGluHieThriHiaHieSleuAaAaAenSerSerLeuSerTrpSerProgly 20
Db 1 ATGGAGCACACGACGCCCTCGCAGCAACAGCTCGCTTGGTGGTCCCGGAC 60
QY 21 SerAlaCyGgLyLeuGlyPheValProValValTYrTYrSerLeuLeuLeuGly 40
Db 61 TCGGCTTCGCGCTTGCGTTTGGTCCGCTGCTTACTACAGGCTTGGTGGTCCGT 120
QY 41 LeuProAlaAenHieLeuThrValHieHieLeuSerGlnLeuValAlaArgArgGlnys 60
Db 121 TTACGACGAATATCTTGAACATGATCTCTCCAGCTGTGGCAAGAGACAGAG 180
QY 61 SerSerTYrAenTYrLeuLeuAlaLeuAlaAaAaAaPheLeuValLeuPhePheIle 80
Db 181 TCTCTCTAACATATCTTCTTGGCAGCTCGCTGCGCAGACATCTGCTCTTTTCA 240
QY 81 ValPheValAaPheLeuLeuGlnAaPheHieLeuAaAaAaAaAaAaAaAaAaAa 240
Db 241 GGTGTGTGATCTTCTGTTGAAGATTTCACTTGAACATGCAATGCTCAGGTCCC 300
QY 101 AapLyIleIleGlnValLeuGlnPheSerSerHieHieHieSerHieTrpIleThrVal 120
Db 301 GACAAATCATGAGATGCTGGAATTCATCATCATCCACCTCCATATGATTACTTA 360
QY 121 ProLeuThrIleAaPArgTYrIleAlaValAlaCySHiSPoleuLYrTHieThrValSer 140
Db 361 CCGTTAACATATGACGAGTAATGCTGTGCGCCCGCTCAAGTACCAACGAGTCCA 420
QY 141 TYrProAlaArgThrArgLYrValIleValSerValTYrIleThrCysPheLeuThrSer 160
Db 421 TACCAAGCCCGACCCGAAAGTCACTTGAAGTGTTCATCAACCTGCTTCTTCAAC 480
QY 161 IleProTYrTYrTrpTrpProAaHieTrpThriGlnAaPheTYrIleSerThrSerValHie 180
Db 481 ATCCCTAATTAAGTGTGCGCCCAATCATGAGTGAAGACTACATCAGACCTCTGTG 540

```

```

QY 181 HieValLeuIleTrpIleHieCysPheThrValTYrLeuValProCysSerIlePhePhe 200
Db 541 CAGTCTCATCTGATTCAGATTCACCTTCCACCGTCTTACCTGTCCTGCTCCATCTTCTTC 600
QY 201 IleLeuAaSerIleIleValTYrIySleuAArgLYrSerAaPheArgLeuArgLY 220
Db 601 ATCTTGAATCATCATCTTGTGTGTAAGCTCAGAGAGAGAGCAATTTTGTCTCGTGGC 660
QY 221 TYrSerThGlyLeuThrThrAlaIleLeuPheThrIleThSerIlePheAlaThrLeu 240
Db 661 TACTCAGCGGAGAGACACCGCCCATCTTGTTCACATTAACCTCCATCTTGGCCACACTT 720
QY 241 TrpAlaProArgIleIleMetIleLeuTYrHieSleuTYrGlyAlaProIleGlnAaArg 260
Db 721 TGGGCCCCCGCATCATCATGATTTTAAACACCTCTATGGGGCCCATCCAGAACGCC 780
QY 261 TrpLeuValHieIleMetSerAaPheIleAlaAaMetLeuAlaLeuAaAaThraIle 280
Db 781 TGGCTGGTGCATCATCATATGTCGACATTCGCAACATGCTAGCCCTTGAACACAGCCATC 840
QY 281 AaenPheLeuTYrCysPheIleSerLYrAaPheArgThrMetAlaAlaThrLeu 300
Db 841 AACTCTTCTCTACTGCTTCTCATACGACGAGTTCGACACCAAGCAGCGCCACGCTC 900
QY 301 LysAlaPhePheLYrCysGlnLYrGlnProValGlnPheTYrThraAaHieAaPheSer 320
Db 901 AAGCTTCTTCAAGTGCAGAGACCAACCTGTATGCTTACACCAATCATTAATCTTTCC 960
QY 321 IleThrSerSerProTrpIleSerProAlaAaSerHieCysIleLYrMetLeuValTYr 340
Db 961 ATTAAGATAGCCCTGATCTGCGCGCAAACTCAACATGCAATCAAGATGCTGTATC 1020
QY 341 GlnTYrAaPheLYrAaGlyLYrProIleLYrValSerPro 353
Db 1021 CAGTATGACAAATGAGAAACCTATAAAGTATCCCCG 1059

```

RESULT 6
ID ACC71785 standard; DNA; 1062 BP.
XX ACC71785;
AC 25-JUL-2003 (first entry)
XX
DE Human G protein-coupled receptor DNA #SEQ ID 1.
XX
KW Human; G protein-coupled receptor; anorectic; eating disorder; obesity;
XX gene; de.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..1062 /*tag= a
FT /*product= "G protein-coupled receptor"
FT
PN W02003027142-A1.
XX
PD 03-APR-2003.
XX
PF 19-SEP-2002; 2002MO-JP009626.
XX
PR 21-SEP-2001; 2001JP-00288278.
XX
PA (YAMA) YAMANOUCHI PHARM CO LTD.
XX
XX Matsumoto S, Takasaki J, Kurama T, Saito T, Kamohara M, Soga T,
PI HiYama H,
XX
DR WPI: 2003-33291/31.
XX
XX P-PSDB; ABR44437.

PT New G protein-coupled receptor and encoded gene, useful in screening
 XX preventives or remedies for eating disorders or obesity.

PS Diacloaure; Page 31-33; 48bp; Japanese.

XX The invention relates to a novel G protein-coupled receptor. The protein
 CC and its encoded gene are useful for screening preventives or remedies for
 CC eating disorders or obesity. The current sequence represents a G protein
 CC coupled receptor encoding sequence

XX Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 other;

Alignment Scores:

Pred. No.:	6 83e-179	length:	1062
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-995-225B-16 (1-353) x ACC71785 (1-1062)

QY 1 MetGluHisThrIleAlaHisLeuAlaAlaAsnSerLeuSerTrpSerProGly 20
 Db 1 ATGGAGACACGCGCCGACCTCGCAGCAGAGCTCGCTGCTGTCGCCCGGC 60
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuGlyLeuGly 40
 Db 61 TCGGCTGCGGCTTGGGTTTCGTGCCGCTGCTACTACAGCCTTTCGCTCGCTCGCT 120
 QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLeu 60
 Db 121 TTTACGACAAATTCCTTGACAGTGAATCATCTCTCCAGCTGGTGGCAAAACAGAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
 Db 181 TCCTCTACACATCATCTTGGCAGCTCGCTGCGGACATCTTGGCTCTTTTCATA 240
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 Db 241 GTGTGTGTGACCTTCTGTTGGAAGATTCATCTTGAAACATGACAGTCTCGAGTCCC 300
 QY 101 AspPheIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 Db 301 GACAAAGATCAAGAAGTCTGGAATTCATCCATCCACCTCCATATGATTAATCTGA 360
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisAspLeuTyrTyrHisThrValSer 140
 Db 361 CGCTTAACCATTAACGATATATCGCTGTCTGCCACCCGCTCAAGTACCAACGCTCTCA 420
 QY 141 TyrProAlaArgThrArgGlyValIleValSerValTyrIleThrCysPheLeuThrSer 160
 Db 421 TACCCAGCCGCGCACCCGGAAGATCATTTGAGTGTAACTACACCTGCTCCGACAGC 480
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrTyrIleSerThrSerValHis 180
 Db 481 ATCCCTATTAATCGTGGCCCAACATCTGAGTGAAGATCAATCAAGCACTCTGTGCAT 540
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 Db 541 CAGCTCTATCTCGATCCACTGCTTCAACGCTTACTGCTGCTGCTGCTCATCTTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrTyrLeuArgGlySerAsnPheArgLeuArgGly 220
 Db 601 ATCTGAATCAATCATTTGTGTCAAGCTCAGAGGAGAGCAATTTTCTCTCCGTCGC 660
 QY 221 TyrSerThrGlySerThrThrAlaIleLeuPheThrIleHisSerIlePheAlaThrLeu 240
 Db 661 TACTCCAGGGGAGAGACACCGCATCTTGTTCACATTAACCTTCATCTTGCACACTT 720
 QY 241 TrpAlaProArgGlyIleIleMetIleLeuTyrHisLeuTyrGlyValProIleGlnAsnArg 260
 Db 721 TGGGCCCCCGCATCATGATTTCTTTACACCTCATGGGGCGCCCATCCAGAACCGC 780

QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 Db 781 TGGCTGTACACATCATGTCGACATTCGACCATCTGAGCTTCTGAAACAGCCATC 840
 QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
 Db 841 AACTTCTTCTCTACTGCTTTCATCCAGCAGCGTTCGACCATGAGCCGCGCAGCTC 900
 QY 301 LysAlaPhePheLeuLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 Db 901 AAGGCTTCTTCTCAAGCCAGCAAGCAACTGTACAGTTCATACCATCATATCTTTTC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleIleMetLeuValTyr 340
 Db 961 ATTAACAAGTACCCCTGGAATCTCGCCGCAACTCACCTGCTCAAGATGCTGTATC 1020
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 Db 1021 CAGTATGACAAATAATGGAATAACTTATTAAGATATCCCG 1059

RESULT 7

ABV73364
 ID ABV73364 standard; DNA; 1202 BP.

XX ABV73364;

XX 22-JAN-2003 (first entry)

XX Human TGR20 polypeptide encoding DNA.

XX G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory;

XX antilipemic; nootropic; neuroprotective; antidiabetic; antidiabetic; human;

XX antiparkinsonian; antilipemic; TGR20; gene; da.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 68..1129

XX FT /tag= a

XX FT /product= "TGR20"

XX FT /note= "GPCR polypeptide"

XX PN MO200277001-A2.

XX PD 03-OCT-2002.

XX PF 08-MAR-2002; 2002MO-US007171.

XX PR 09-MAR-2001; 2001US-00802803.

XX PR 16-MAR-2001; 2001US-0276649P.

XX PA (TUL-) TULARIK INC.

XX PI Tian H, Zhao J, Chen J, Cutler G;

XX DR WPI, 2003-01881/01.

XX DR P-PSDB; ABB82499.

XX PT New G-protein coupled receptor polypeptides and polynucleotides useful

XX for identifying compounds for treating a TGR-associated disorder, e.g.

XX psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's

XX disease, anemia.

XX Claim 6; Page 63; 87pp; English.

XX The invention relates to G-protein coupled receptor (GPCR) polypeptides

XX and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR83,

XX TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are

XX useful for identifying compounds for treating a TGR-associated disorder,

XX such as psoriasis, inflammatory bowel disease, hyperlipidemia,

XX Parkinson's disease, Huntington's disease, anemia, immune and blood

XX disorders, ulcerative colitis, Crohn's disease or spleen enlargement.

CC They are also useful for identifying cells such as kidney, liver,
 CC hypothalamus, colon, adipose, or spleen cells, for forensics and
 CC paternity determination, diagnosing diseases and examining signal
 CC transduction. The present sequence represents a human TOR20 polypeptide
 CC encoding DNA
 XX

Sequence 1202 BP; 267 A; 380 C; 258 G; 297 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8,1e-179	Length:	1202
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-09-995-225b-16 (1-353) x ABV73364 (1-1202)

QY 1 MetGluHisThrHisAlaHisAlaAsnSerLeuSerTrpSerProGly 20
 DB 68 ATGAGAGACACGACGCCCTCGACGACCAAGCTCGCTTGTTGGTGTCCCGCG 127
 QY 21 SerAlaCyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCyGly 40
 DB 128 TCGGCTGGGGCTTGGGTTTGGCCGTGCTACTACAGCCTCTTGGCTCGGT 187
 QY 41 LeuProAlaAsnLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGln 60
 DB 188 TTACACGACAAATCTTGACAGATCATCTCTCCAGCTGTGGCAAGAACAGAG 247
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
 DB 248 TCCTCTACAACTATCTTGGCAGCTGCTGCGACGACATCTTGCTCTTTTCATA 307
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetProGlnValPro 100
 DB 308 GTTGTGTGACTTCTGTGGAAGATTCATCTTGAAATGACAGATGCTCAGGTCCC 367
 QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 368 GACAAATCATGAAGTGTGGAAATCTCATCATCACTCCATAGGATTACTGTA 427
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
 DB 428 CCGTTAACATGACGATATCGCTGTGCGCCCGCTCAAGTACACACGAGCTCA 487
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 488 TACCCAGCCCGACCCGGAAGTCAATTTGAAGTTTACATCACTGCTTCTGACCAAG 547
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpTrpGluAspTyrIleSerThrSerValHis 180
 DB 548 ATCCCTATTAATCGGTGGCCCAACATCTGACAGAACTAATACAGCACTCTGTGAT 607
 QY 181 HisValIleIleIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 608 CACGCTCATCTGATTCACATGCTTCACCGCTCTGCTGCTGCTGCTGCTGCTGCT 667
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgGlySerAsnPheArgLeuArgGly 220
 DB 668 ATCTTGAACTCAATCTTGTATCAAGCTCAGAGAGAGAGAAATTTGTTCCGTGGC 727
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 728 TACTCACGGGAGAACACCGCCCATCTTGTTCACCATTACTCATCTTGGCCACACT 787
 QY 241 TrpAlaProArgGlyIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 788 TGGGCCCCCGCATCATATGATTTTAAACCACTTATGGGGCCCGCATCAGAAACGCG 847
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 848 TGGCTGTATGACATATGTCGACATTTGCCAATGCTAAGCCCTTTCTGAACACAGCCATC 907

QY 281 AsnPhePheLeuTyrCysPheIleSerIleValArgPheArgThrMetAlaAlaThrLeu 300
 DB 908 AACCTTCTCTTACTGCTTCAATCAGACCGGTTCCGACCATGACGCCCGCACGCTTC 967
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 968 AAGCTTTCTTCAAGTGCAGAGAACCTTGACGTTCTTACACCAATCATATCTTTCC 1027
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 DB 1028 ATAAACAGTAGCCCTCGATCTCGCCGCAACTCACACTGCATCAAGTGTGTTGAC 1087
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1088 CAGTATGACAAATAATGAAAAACCTATAAAGTATCCCG 1126

RESULT 8
 ADF70587
 ID ADF70587 standard; DNA; 1776 BP.
 XX
 AC ADF70587;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Orphan receptor ligand-related human protein gene SeqID210.
 XX
 KW ligand; orphan receptor protein; fusion protein; fluorescent protein;
 KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
 KW GFPuv; Enhanced GFP; EGFP; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003071272-A1.
 XX
 PD 28-AUG-2003.
 XX
 PF 21-FEB-2003; 2003WO-JP001901.
 XX
 PR 22-FEB-2002; 2002JP-00045728.
 PR 23-JUL-2002; 2002JP-00213949.
 PR 11-OCT-2002; 2002JP-00298237.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
 XX
 DR WPI; 2003-697654/66.
 DR P-PDB; ADF70485.
 XX
 PT Transformation of cells with a fusion protein of an orphan receptor
 PT protein with a fluorescent protein useful for identification of ligands
 PT to the orphan receptor.
 XX
 PS Example 4; SEQ ID NO 210; 594bp; Japanese.
 XX
 CC This invention relates to a novel method of identifying ligands to an
 CC orphan receptor protein which comprises transforming cells with DNA
 CC encoding a fusion protein of the orphan receptor with a fluorescent
 CC protein, so that the fusion protein is expressed in the cells (or cell
 CC membranes isolated from them) and contacting the cells with the potential
 CC ligand to be tested. A suitable fluorescent protein (GFP) for incorporation in
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
 CC identification of ligands binding to an orphan receptor protein.
 XX

Sequence 1776 BP; 471 A; 500 C; 358 G; 447 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,39e-178	Length:	1776
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0
US-09-995-225B-16 (1-353) x ADF70587 (1-1776)			

OY	1	I	NetG1uH1	gTHrH1eAlaH1s1eLeuAla1aAaNsSerSerLeuSerTrpTrpSerProGly	20
Db	1	A	TGAGGACACAGCAGCCACCCTGGCACCAACAGCTGCCTGTCTTGGTGTCTCCCGGC	60	
OY	21	S	erAlaCyAg1yLeuG1yPheValProValValTyrTyrSerLeuLeuCybLeuG1y	40	
Db	61	T	CGGCGTGGGCTTGGGTTTGGTGGCCCTGGTCTACTACAGCTTCTGTCTGCTCGGT	120	
OY	41	L	euProAlaam11eLeuThrVal11e11eLeuSerG1uLeuVal1a1aArgSG1u1yS	60	
Db	121	T	TACACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAG	180	
OY	61	S	erSerTyrAnt1yTyrLeuLeuAla1a1a1aAaP11eLeuValLeuPhePhe11e	80	
Db	181	T	CTCTCTTACAACTATCTTGGCACTGGCTGCGCCAGCAATCTGGTCTCTTTTTCATA	240	
OY	81	V	alPheValaAaPheLeuLeuG1uAaPhe11eLeuAaMetG1uMetProG1uValPro	100	
Db	241	G	TGTTTGTGGACCTTCTCTGTGGAGATTTCATCTTGAACATGGCAATGCTCAAGTCCC	300	
OY	101	A	sp1yS11e11eG1uVal11eG1u1uPheSerSer11eH1srHSer11eTrp11eThrVal	120	
Db	301	G	ACAAAGATCAATAGAAGTCTGGAAATTCATTCATCCACACCTCATATGATTAATGTA	360	
OY	121	P	roLeuThr11eAaPArgTyr11eAlaValCyH1sProLeu1yS1yH1sThValSer	140	
Db	361	C	GGTAAACCATTTGACAGGTATATGCTGTCTGGCCACCGGCTCAAGTACACACAGGTTCA	420	
OY	141	T	yrProAlaArgThrArg1yVal11eValSerValTyr11eThrCybPheLeuThSer	160	
Db	421	T	ACCAGGCGCGACCCGGAAAGTCAATTGTAAAGTTTACATCACTGCTTCTACACAGC	480	
OY	161	I	11eProTyrTyrTrpTrpProAaM11eTrpThrg1uAaP1yTyr11eSerThSerValH1s	180	
Db	481	A	TCCCTATATCTGGTGGCCCAACATCTGGACGTAAAGCTACATCAGCAACCTCTGTGAT	540	
OY	181	H	1sVal11eLeu11eTrp11eH1sCybPheThrValTyr11eLeuValProCySer11ePhePhe	200	
Db	541	C	AGCTCTCATCTGGATCATCTGCTTCCACCGTCTACTGCTGGCCCTGCTCATCTTCTTC	600	
OY	201	I	11eLeuAaSer11e11eValTyr1y1yLeuArgArg1ySserAaPheArg1eAaArg1y	220	
Db	601	A	ATCTGAACCTCAATCATTTGTGTACAGCTCAGAGGAAGACAAATTTTCGCTCGTGGC	660	
OY	221	T	yrSerThrArg1yS1yThrThrAla11eLeuPheThr11eThrSer11ePheAlaTh1eu	240	
Db	661	T	ACTTCACAGGGGAAGACACCGCAATCTGTTCACCATTAATCTTCATTTGGCACACTT	720	
OY	241	T	ryAlaProArg11e11eMet11eLeuTyrH1sLeuTyrG1yAlaPro11eG1uAaAArg	260	
Db	721	T	GGGCCCCCGCATCATCATATCTTTTACCAACCTCTATGGGGGCCCATTCAGAACGCG	780	
OY	261	T	rpLeuValH1a11eMetSerAaP11eAlaAaMetLeuAlaLeuLeuAaThrAla11e	280	
Db	781	T	GGCTGGGACATCATATGCTCCGACATTCGCCAAATGTAGGCCCTTGTAAACAGACCATC	840	
OY	281	A	asPhePheLeuTyrCybPhe11eSer1yAaArgPheAaG1uThrMetAla1a1aTh1eu	300	
Db	841	A	ACTTCTTCTCTACTGCTTATACAGCAAGGAGTTCGACACATGGCAGCGCCACAGCTC	900	
OY	301	L	yAla1aPhePhe1ySCySG1u1yG1uProValG1uPheTyrThrAaM1aAaPheSer	320	
Db	901	A	AGGCTTTCTTCAAGTGCAGAGAACTGTACAGTTCTTACCAATCATTAATTTTCC	960	
OY	321	I	11eThrSerSerProTrp11eSerProAlaAaSerH1sCyb11e1ySMeLeuValTyr	340	
Db	961	A	TAAACAATAGACCCCTGGATCTCTGGCGGCAAACTACACTGCATCAAAATCTGGTGTAC	1020	

QY	341	GlnTyrAspIysAsnGlyLysProIleLeuValSerPro	353
Db	1021	CAGTATGACAAATAATGGAAAACCTATAAAGTATCCCCG	1059

RESULT 9	
AAAS15731	AAAS15731 standard; cDNA, 1343 BP.
ID	AAAS15731
AC	AAAS15731;
XX	
DT	14-FEB-2002 (first entry)
XX	
DE	DNA encoding chemokine receptor family related protein, NOV11.
XX	
XX	NOV; cytosolic; psoriasis; neutropenic; neuroprotectant; cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia; haemostatic; atherosclerosis; gene therapy; neurogenesis; motility; differentiation; proliferation; hematopoiesis; wound healing; angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis; angiodysplasia; allergy; pendred syndrome; skeletal dysplasia; ischaemic injury; neuroepithelial disorder; hepatitis; heart failure; chemokine receptor; chromosome 1; ss.
KW	
KW	Homo sapiens.
XX	
OS	
XX	
Key	Location/Qualifiers
FT	1
FT	5'UTR
FT	/*tag= a
FT	2..1063
FT	/*tag= b
FT	/*product= "Chemokine receptor related protein, NOV11"
FT	2..142
FT	/*tag= c
FT	143..1060
FT	/*tag= d
FT	/note= "Mature chemokine receptor related protein, NOV11"
FT	1064..1343
FT	/*tag= e
FT	
XX	
XX	MO200170978-A2.
PN	
XX	
PD	27-SEP-2001.
XX	
PE	20-MAR-2001; 2001MO-US009093.
XX	
PR	20-MAR-2001; 2000US-0190768P.
PR	20-MAR-2001; 2000US-0190835P.
PR	22-MAR-2001; 2000US-0190972P.
PR	22-MAR-2001; 2000US-0191199P.
PR	24-MAR-2001; 2000US-0191947P.
PR	28-MAR-2001; 2000US-0192857P.
PR	28-MAR-2001; 2000US-0192664P.
PR	28-MAR-2001; 2000US-0192665P.
PR	28-MAR-2001; 2000US-0192984P.
PR	29-MAR-2001; 2000US-0192836P.
PR	31-MAR-2001; 2000US-0193843P.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
XX	Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;
PI	Vernet CM;
XX	
DR	WPI, 2001-639127/73.
XX	
DR	P-PSDB; AA010068.
XX	
PT	Polypeptides and nucleic acids related to chloride channel, insulin-like growth factor family of proteins, useful for diagnosing and treating cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
XX	
PS	Claim 9, Page 44; 15ipp; English.
XX	
XX	The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX polypeptides are useful for treating pathology associated with NOVX

CC polypeptide, determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX, identifying agents binding to
 CC NOVX and treatment of disorders associated with altered expression of
 CC members of chloride channel-associated proteins e.g. cystic fibrosis and
 CC congenital myotonia. NOVX proteins are useful in treatment of disorders
 CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,
 CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
 CC in one or more organs (e.g. haemophilia, anaemia), pendred syndrome,
 CC skeletal dysplasias, disorders characterised by altered cell shape,
 CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial
 CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
 CC treatment of disorders of vascular smooth muscle cell differentiation.
 CC e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are
 CC useful to screen for molecules which inhibit or enhance NOVX activity or
 CC function and are useful as targets for the identifying small molecules,
 CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
 CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV
 CC sequences are also useful for: identifying a cell or tissue type in a
 CC biological sample; amplifying DNA sequences from very small biological
 CC samples e.g. hair or skin or body fluids and as primers and probes to
 CC identify and/or clone NOVX homologues. NOVX proteins are useful
 CC immunogens to generate antibodies to monitor protein levels and modulate
 CC NOVX activity. Cells comprising the nucleic acids are useful for
 CC producing transgenic animals, for studying the function and/or activity
 CC of NOVX protein and identifying and/or evaluating modulators of NOVX
 CC protein activity. This sequence encodes NOV11 (located on chromosome 1)
 CC related to the chemokine receptor family of proteins, one of 12 NOV
 CC polypeptides described in the method of the invention
 XX

SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,39e-178 Length: 1343
 Score: 1868.00 Matches: 352
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 1
 Query Match: 99.79% Indels: 0
 DB: 4 Gaps: 0

US-09-995-225b-16 (1-353) x AAS15731 (1-1343)

QY 1 MetGLHIETHRHISALHISLEUALAIAENSSerSerLeuSerTrpTrpSerProgly 20
 DB 2 ATGAGACACACGACGCCACCTCGACGACCAACAGCTCGCTGTGGTGCGCCCGCGC 61
 QY 21 SerAlaCyEGlyLeuGlyPheValProValValIYTYTSerSerLeuLeuGly 40
 DB 62 TCGGCTCGCGGCTTGGGTTTCGGCCCGTGGTCTACTACAGCTCTTGGTCTCGGT 121
 QY 41 LeuProAlaAenIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnlys 60
 DB 122 TTACACGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGTGGCAAGAGACAGAG 181
 QY 61 SerSerTYASnTYRLeuLeuAlaLeuAlaAlaIleAaPIIleLeuValLeuPhehelle 80
 DB 182 TCCTCTCAACAATATCTCTTGGCATCTCGCTGCGCAATCTTGGCTCTTCTTCTTA 241
 QY 81 ValPheValAspPheLeuLeuGlnAspPheIleLeuAaMetGlnMetProGlnValPro 100
 DB 242 GTGTTTGGAGCTCTCTGTGGAGATTTCATCTTGAACATGCAATGCTCAGGTCCC 301
 QY 101 AspLysIleIleGlnValLeuGlnPheSerSerIleHISTrpSerIleTrpIleThrVal 120
 DB 302 GACAAAGATCATAGAGTCTGGAATTCATCATCCACACCTCATATGATTAAGTGA 361
 QY 121 ProLeuThrIleAspArgTYRILEalValCysHISProLeuLysTYRHISThrValSer 140
 DB 362 CCGTTAACCATTTGACAGGATATATACGTCTGCCACCGCTCAAGTACCAAGGATCA 421
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTYRIleThrCysPheLeuThrSer 160
 DB 422 TACCAAGCCCGACCCGGAAGATGATGTAGTGTATCATCACTGCTTCTTCAACAGC 481

QY 161 IleProTYRTrpTrpProAenIleTrpTrpGlnAspTYRILESerThrSerValHis 180
 DB 482 ATCCCTATTACTGCGTGGGCCCAACATCTGAGCTGAAGACATACAGCACTCTGTGAT 541
 QY 181 HisValIleIleTrpIleHISCySPheThrValTYRLeuValProCySerIlePhePhe 200
 DB 542 CAGCTCTCATCTGATTCACAGCTTTCACCGCTCACTGAGTGCCTGTCCATCTTCTTC 601
 QY 201 IleLeuAaSerIleIleValTYRLeuAaArgLysSerAaPheArgLeuArgGly 220
 DB 602 ACTTGAATCATATCATTTGTGACAGCTCAAGAGAAAGACAAATTTGCTCCGTGAC 661
 QY 221 TYRSerTrpGlyLysSerThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 662 TACTCCAGCGGGAAGACACCGGCATCTTGTTCACCATTAATCTCATCTTTGCCACATT 721
 QY 241 TrpAlaProArgIleIleMetIleLeuTYRHISLeuTYRGLValAlaProIleGlnAaArg 260
 DB 722 TGGGCCCCCGCATCATATGATTTCTTACCACTCTATGAGGCGCCCATCCAGAACCGC 781
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAaMetLeuAlaLeuAaThrAlaIle 280
 DB 782 TGGCTGTACATCATATGTCGACATTCGCAACATGCTAGCCCTTCTGACACAGCCATC 841
 QY 281 AaPhePheLeuTYRCysPheIleSerIysArgPheArgThrMetAlaIleAlaThrLeu 300
 DB 842 AACTCTTCTCTCATCTGCTTCATCAGCAAGCGGTTCGCAACATGAGCGCCGACGCTC 901
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTYRThrAaHisAaPheSer 320
 DB 902 AAGCTTTCCTTCAAGTCCGACAGACCAACCTGTACAGTTCTACCAATCAATCACTTTTC 961
 QY 321 IleThrSerSerProTrpIleSerProAlaAaSerHisCysIleLysMetLeuValTYR 340
 DB 962 ATTAACAAGTACCCCTGGAATCTCGCGGCAAACTCAACATGACATCAAGATGCTGTATC 1021
 QY 341 GlnTYRAspLysAaGlyLysProIleLysValSerPro 353
 DB 1022 CAGTATGCAAAATATGAAAAACCTATTAAGATATCCCG 1060

RESULT 10
 ADJ87768
 ID ADJ87768 standard; DNA, 1343 BP.
 XX
 AC ADJ87768,
 DT 06-MAY-2004 (first entry)
 XX
 DE G-coupled protein receptor-related protein coding sequence #57.
 KW novel protein; G-coupled protein receptor-related protein;
 KW cardiomyopathy; atherosclerosis; cell signal processing-related disorder;
 KW metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
 KW Huntington's disease; epilepsy; anxiety; pain; hypercholesterolemia;
 KW obesity; hypertension; Crohn's disease; systemic lupus erythematosus;
 KW viral infections; bacterial infection; parasitic infection;
 KW hyperthyroidism; hypothyroidism; von Hippel-Lindau syndrome;
 KW Alzheimer's disease; tuberosus sclerosis; hypercalcaemia; cerebral palsy;
 KW gene, ds.
 OS Unidentified.
 PN WO2002102321-A2.
 XX
 PD 27-DEC-2002.
 PF 18-JUN-2002; 2002WO-US019522.
 XX
 PR 18-JUN-2001; 2001US-0298994P.
 PR 18-JUN-2001; 2001US-0299134P.
 PR 04-OCT-2001; 2001US-00972446.
 PR 06-JUN-2002; 2002US-00291134.
 PR 07-JUN-2002; 2002US-00298994.

XX (CURA-) CURAGEN CORP.
 XX
 PA Anderson DW, Guo X, Gusev VY, Hermann JL, Li L, Mezes PS;
 PI Pena CE, Spaderna SK, Zhong M;
 XX WPI; 2003-167441/16.
 DR P-PSDB; ADJ87769.
 XX
 PT New MOLX polypeptides and polynucleotides, useful in gene therapy,
 PT particularly for treating or preventing e.g. cardiomyopathy,
 PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
 PT cancer, stroke or pain.
 PS
 PS Claim 8; SEQ ID NO 203; 378bp; English.
 XX
 XX The invention comprises the amino acid and coding sequences of novel G-
 CC coupled protein receptor-related (MOL) proteins. The DNA and protein
 CC sequences of the invention are useful for treating or preventing a MOL-
 CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
 CC associated with cell signal processing and metabolic pathway modulation,
 CC or diabetes. The DNA and protein sequences are also useful for the
 CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
 CC pain, hypercholesterolemia, obesity, hypertension, Crohn's disease,
 CC systemic lupus erythematosus, viral infections, bacterial infections,
 CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
 CC syndrome, Alzheimer's disease, tuberous sclerosis, hypercalcaemia, or
 CC cerebral palsy. The present DNA sequence encodes a MOL protein of the
 CC invention.
 XX
 SO Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2.39e-178 Length: 1343
 Score: 1868.00 Matches: 352
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 1
 Query Match: 99.79% Indels: 0
 DB: 10 Gaps: 0
 US-09-995-225b-16 (1-353) x ADJ87768 (1-1343)
 QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpSerProGly 20
 DB 2 ATGGAGCAGCAGCAGCCGACCTGCGAGCAAGCGCTGCTGGTGTGCCCCGGC 61
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
 DB 62 TCGGCTGGCGGCTTGCGTTTGTGGCCGCTGTCTACACGCTCTTGGTGTGCTCGGT 121
 QY 41 LeuProAlaAsnLeuThrValAlaLeuSerGlnLeuValAlaAlaGArgGlnLys 60
 DB 122 TTACGAGCAAAATCTTGACAGGATCATCTCTCCAGCTGGTGGCAAGACAGAG 181
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspLeuLeuValLeuPhePhe 80
 DB 182 TCTCTCCACACATCTCTTGGGACCTGCTGCGCGACACTTGGCTCTTTTTCATA 241
 QY 81 ValPheValAspPheLeuLeuGlyLeuAspPheLeuLeuAsnMetGlnMetProGlnValPro 100
 DB 242 GTGTTTGTGACCTTCTGTTGGAAAGATTCTTGAACATGACAGATGCTCAGTCC 301
 QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 302 GACAAACATCAAGAGTGTGGAAATCTCATCATCATCACACCTCCATATGATTACTGTA 361
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
 DB 362 CGCTTAACCTTAACAGATATACACTGTCTGCCACCCGCTCAAGATCCACAGGCTCTCA 421
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 422 TACCGAGCCGCAACCGGAAGTCAATTGTAAGTTTATCATCACTGCTTCTGAGCAGC 481

QY 161 IleProTyrTyrTrpTrpProAlaSerIleTrpThrGluAspTyrIleSerHisSerValHis 180
 DB 482 ATCCCTATTACTAGGAGCCCAACATCTGACAGAAAGACTACATTCGACCTTGAGCAT 541
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 542 CAGCTCTCATCTGATTCACATGCTTACCGCTTACCTGTGCTCCGCTCCATCTTCTTC 601
 QY 201 IleLeuAsnSerIleIleValTyrIysLeuArgArgLysSerAsnPheAlaArgLeuArgGly 220
 DB 602 ACTTGAACTCAATCATCTGTGTAAAGCTCAGAGAGAGAGCAATTTTGCTCCGTGCG 661
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 662 TACTCCACGGGAGAGCCAGCCCATCTTGTTCACCATTAATCTCCATCTTGGCACACTT 721
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 722 TGGGCCCCCGCATCATCATGATTTCTTACACCTTATGGGCGCCCATCCAGAACCGC 781
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 782 TGGCTGTGACATCATGTGCCACATGSCCAATGCTAGCCCTTGAAACAGCCATC 841
 QY 281 AsnPhePheLeuTyrCysPheIleIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
 DB 842 AACTCTTCTCTTACGCTTCAATCAGCAAGCGGTTCGCAATGGCACCGCCAGCGTC 901
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 902 AAGCTTTTCTTCAAGGCCAGAAAGCACTTGACAGTTTACACCAATATATCTTTTC 961
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 DB 962 ATPACAGAGAGCCCGCTGATCTCGCCGCAACTCACACTGCATCAAGATGCTGCTGAC 1021
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1022 CAGTATGACAAATAAGAAACCTATTAAGATATCCCG 1060
 RESULT 11
 AD179324
 ID AD179324 standard; DNA; 1343 BP.
 XX
 AC AD179324;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE NOV11 coding sequence, SEQ ID 21.
 XX
 XX Cytostatic; Immunosuppressive; Antiallergic; Antimicrobial; Vasotropic;
 XX Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;
 XX Ophthalmological; Antipneumatic; Neuroprotective; Nootropic;
 XX Antiartherosclerotic; Hypotensive; Cardiant; Cerebroprotective;
 XX Gene Therapy; NOVA; human; cancer; myelogenous leukaemia;
 XX congenital neonatal autoimmune chromocytopenia; immunological disorder;
 XX allergy; infection; asthma; lung disease; reproductive disorder;
 XX haemangoma; deafness; liver cirrhosis; hepatitis C; gastric disorder;
 XX diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;
 XX hypertension; stroke; heart failure; chromosome 1; NOV11;
 XX chemokins receptor; gene; de.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 2..1063
 FT /tag= a
 FT /product= "NOV11"
 XX
 XX MO2004009635-A2.
 XX
 XX 29-JAN-2004.

XX 04-OCT-2001; 2001WO-US031292.
PF
XX 20-MAR-2001; 2001US-00813432.
PR
XX (CURA-) CURAGEN CORP.
PA
XX Taupier RJ, Majumder K, Spaderna SK, Smtneon G, Mezes PS;
PI Verzet CM;
XX WPI, 2004-123380/12.
DR P-PSDB; ADI79325.
XX
PS Claim 8; Page 43-44; 158bp; English.
XX
CC The present invention relates to novel NOVX proteins and their coding
CC sequences (ADI79304-ADI79327). The sequences are useful for the
CC manufacture of a medicament for treating a syndrome associated with a
CC disorder associated with the NOVX protein, or for diagnosing and treating
CC disorders associated with the NOVX protein, such as cancer, myelogenous
CC leukaemia, congenital neonatal autoimmune thrombocytopenia,
CC immunological disorders, allergy and infection, asthma, lung diseases,
CC reproductive disorders, male and female reproductive diseases,
CC hemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,
CC diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,
CC hypertension, stroke and heart failure. NOV1 represents a new member of
CC the chemokine receptor family and the gene is located on chromosome 1.
XX
SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,398-178 Length: 1343
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 12 Gaps: 0
US-09-995-225B-16 (1-353) x ADI79324 (1-1343)
QY 1 MetGLuHfIeThRiSaIaHfIeUaIaAaNsSerIeUeSerTrpTrpSerProGly 20
DB 2 ATGAGACACACGACCCGACCTCGACGCAAGCTGCTGTTGGTGGCCCCGGCC 61
QY 21 SerAlaCySgLYLeuGlyPheValProValIYrYrSerIeUeUeUeCysLeuGly 40
DB 62 TCGGCTGGCGGCTTGGGTTTCGCGCCGTTGCTACTACAGCTCTGCTGCTCGT 121
QY 41 LeuProAlaAsnIleUeUeThrValIleIleUeSerGlnLeuValAlaArgArgGlnys 60
DB 122 TTAACGCAAAATATCTTGACAGTGCATCTCTCCACCTGCTGCGCAAGACAGAA 181
QY 61 SerSerTYrAenTYrLeuUeUaIaUeUaIaAaNsSerIeUeUeUeUeUeUeUe 80
DB 182 TCCTCTCAACATATCTTCTGCACTCGCTGCTGCGACATTTTGCTCTTTTCA 241
QY 81 ValPheValAspPheUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUe 100
DB 242 GTGTTGTGACCTCTGTTGGAAGATTCATCTTGACATGCAATGCTCAGGTCCC 301
QY 101 AApLYsIleIleGlnValLeuGlnUeUeSerIleHfIeThSerIleTrpIleThVal 120
DB 302 GACAGATCATTAAGTGTGAAATTCATTCATCCACACCTCATATGATTAAGTCT 361
QY 121 ProLeuThrIleAspArgTYrIleAlaValCysHfIeProLeuTYrThfIeThValSer 140
DB 362 CCGTTAACCATTCAGAGATATCACTGTCTGCCACCGCTCACTACACAGGCTTCA 421
QY 141 TYrProAlaArgThfIeThArgValIleValSerValTYrIleThCysPheUeThSer 160
DB 422 TACCCAGCCCGACCCGAAAGTCAATGATGTTTACATCACTGCTTCTGACGAGC 481
QY 161 IleProTYrTYrTrpTrpProAenIleTrpThrGlnAepTYrIleSerThSerValHis 180

DB 482 ATCCCTATTACTGAGTGGCCCAACATCTGACTGAAGACTATCATGACGACTCTGTCAT 541
QY 181 HfIeSvalLeuIleTrpIleHfIeSgPheThrValTYrLeuValProCysSerIlePhePhe 200
DB 542 CAGTCTCTATCTGAGTCCACGCTTCACCGCTTCACCGTTCCTGCTCCATCTTCTTC 601
QY 201 IleUeUeSerIleIleValTYrLYsLeuAArgLYsSerAspPheArgLeuArgLY 220
DB 602 ATCTGAACATCATATGTTGTACAGCTCAGAGAGAGAGCAATTTTGCTCCGTGCC 661
QY 221 TYrSerThrGlyLYeThrThfIeAlaIleUeUeThrIleThSerIlePheAlaThrLeu 240
DB 662 TACTCCAGCGGAGAACACACCGCATCTTGTTCACCATTAACCTTCATCTTCCACACTT 721
QY 241 TrpAlaProArgIleIleUeUeThrIleUeUeThfIeUeUeUeUeUeUeUeUeUe 260
DB 722 TGGGCCCCCGCATATCATATCTTTTACCACTCTTGGGCGCCCATCCAGAACCGC 781
QY 261 TrpLeuValHfIeIleMetSerAspIleAlaAsnMetLeuAlaLeuUeUeThrAlaIle 280
DB 782 TGGCTGTACACATCATGTCCGACATTCGCAACATGCTAGCCCTTGACACAGCCATC 841
QY 281 AsnPheUeUeTYrCysPheIleSerLYsArgPheArgThrMetAlaAlaThrLeu 300
DB 842 AACCTCTCTCTACTGCTTCAATCAGACGCGGTTCCGACCATGACGCGCCGCTCC 901
QY 301 LYsAlaPhePheLYsCYeGlnLYeGlnProValGlnPheTYrThrAsnHfIeAspPheSer 320
DB 902 AAGCTTTCTTCAAGTGCAGAGCAAGCACTGTACAGTTCTACACCAATCATTAACCTTTCC 961
QY 321 IleTrpSerSerProTrpIleSerProAlaAsnSerHfIeCYsIleUeMetLeuValTYr 340
DB 962 ATAAACATACGCCCTGATCTCGCCGCAAACTCACATGCAATCAAGATGCTGTATC 1021
QY 341 GlnTYrAspLYsAsnGlyLYeProIleUeValSerPro 353
DB 1022 CAGTATGCAAAATGAGAAACCTTAATAAGTATCCCG 1060
RESULT 12
ADO56003
ID ADO56003 standard; cDNA; 1343 BP.
XX
AC ADO56003;
XX
PT 15-JUL-2004 (first entry)
XX
DE DNA encoding human NOV1.
XX
KW human; gene; ss; cancer; obesity; diabetic nephropathy;
KW acute pancreatitis; stroke; multiple sclerosis.
XX
OS Homo sapiens.
XX
PD US2004058862-A1.
XX
PD 25-MAR-2004.
XX
PD 18-SEP-2002; 2002US-00246583.
XX
PR 18-SEP-2002; 2002US-00246583.
XX
PA (MAJU/) MAJUMDER K.
XX
PI Majumder K;
XX
DR WPI, 2004-26835/25.
DR P-PSDB; ADO56004.
PT Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic
PT nephropathy, acute pancreatitis, strokes and multiple sclerosis.
XX
PS Disclosure; Page 28; 87bp; English.

XX The invention relates to novel isolated NOVX nucleic acids and encoded
 CC polypeptides. The nucleic acids, polypeptides and antibodies raised
 CC against the polypeptides are useful for preventing or treating diseases
 CC associated with aberrant NOVX expression or activity e.g., cancer,
 CC obesity, diabetic neuropathy, acute pancreatitis, stroke, multiple
 CC sclerosis. The present sequence represents a NOVX nucleic acid of the
 CC invention.

XX Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,39e-178	Length:	1343
Score:	1868.00	Matches:	352
Percent Similarity:	99.72%	Conservative:	0
Best Local Similarity:	99.72%	Mismatches:	1
Query Match:	99.79%	Indels:	0
DB:	12	Gaps:	0

US-09-995-225b-16 (1-353) X ADO56003 (1-1343)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpTrpSerProGly 20
 DB 2 ATGAGACACACGACCGCCACTCGCAGCCAGCTCGCTTGTGTGTCCTCCCGC 61
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuGly 40
 DB 62 TCGGCGCTGGCTGGTTCGTGCGCTGCTACACAGCTTGTGCTGCTGCTGCT 121
 QY 41 LeuProAlaAsnLeuThrValIleLeuSerGlnLeuValAlaArgArgGlnLys 60
 DB 122 TTACACGAAATCTTGAAGATGATCTCTCCAGCTGGTGGCAAGACAGAG 181
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
 DB 182 TCCCTCAACAATATCTCTGGCAGCTGCTGCGCAGACATCTTGGCTCTTTTCATA 241
 QY 81 ValPheValAspPheLeuLeuGlnPhePheIleLeuAsnMetGlnMetProGlnValPro 100
 DB 242 GTTGTGTGACCTTCTGTTGGAAGATTATCTTGAACATGACAGATCGCTGCTCC 301
 QY 101 AspLeuIleIleGluValLeuGlnPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 302 GACAAAGTCTAAGAGTCTGAGATTCTCATTCACACCTCCATATGATTAATCTGTA 361
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuTyrTyrHisThrValSer 140
 DB 362 CCGTTAACATATGACAGTATATCACTGCTGCGACCCGCTCAAGTACACACGGTCTCA 421
 QY 141 TyrProAlaArgThrArgValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 422 TACCCAGCCCGCACCCGGAAGATCATTTGTAAGTTCATCACTGCTTCTGACAGC 481
 QY 161 IleProTyrTyrTrpProAsnIleTrpThrGlnAspTyrIleSerThrSerValHis 180
 DB 482 ATCCCTATTAATCTGAGTGGCCCAACATTTGACATGAGACTACATCAGACCTCTGTCAT 541
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 542 CACGTCCTCATCTGATCCACCTGCTTACCGTCACTGCTGCTGCTCATCTTCTTC 601
 QY 201 IleLeuAsnSerIleIleValTyrTyrLeuArgArgLysSerAsnPheArgLeuArgGly 220
 DB 602 ATCTTGAACATCAATGTGTGACAGCTGAGAGGAGCAATTTTCTCTCCGTCGC 661
 QY 221 TyrSerThrGlyTyrThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 662 TACTCAGCGGGAAGACACCGCATCTTTCACCACTTACCTCCATCTTTCACACTT 721
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 722 TGGGCCCCCGCATCATGATTTCTTTACACCTGATGAGGCGCCATCCAGAACCGC 781

QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 782 TGGCTGTACACATCAATGTCGACATTCGCAACATCATCTGACCTTTCGAAACAGCCATC 841
 QY 281 AsnPhePheLeuTyrCysPheIleSerIleArgPheArgThrMetAlaAlaThrLeu 300
 DB 842 AACTTCTTCTCTACAGCTTCATCAGCAAGCGGTTCGCAACATGACAGCCGACGCTC 901
 QY 301 LysAlaPhePheLeuCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 902 AAGCTTCTTTCAGAGTCCAGAGCAACCTGTACAGTTCTACACATATATCTTTC 961
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 DB 962 ATACAGAGTACCCCTGAGATCTGCGGCAAACTCACATGCAATCAAGATGCTGTGTC 1021
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1022 CAGTATGACAAATAATGAAAAAAGCTTAAAGATATCCCG 1060

RESULT 13

ADO28955 standard; cDNA; 2273 BP.

ADO28955;

29-JUL-2004 (first entry)

Human novel GPCR PGR3 polynucleotide, SEQ ID NO:54.

G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; anti-nausea;
 KW cytostatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
 KW CNS; central nervous system; respiratory; antidiabetic; antidiabetic;
 KW virulide; hepatotropic; antibacterial; antianemic; antidiabetic;
 KW dermatological; antitumor; antihypertensive; antiallergic; anorectic;
 KW immunosuppressive; nephrotoxic; gene therapy; GPCR modulator; human;
 KW gene; ss.
 KW Homo sapiens.
 KW OS
 KW PN WO2004040000-A2.
 KW PD 13-MAY-2004.
 KW PP 09-SEP-2003; 2003WO-US028226.
 KW PR 09-SEP-2002; 2002US-0409303P.
 KW PR 09-APR-2003; 2003US-0461329P.
 KW PA (PRIM-) PRIMAL INC.
 KW PI Galienaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
 KW PI Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
 KW DR WPI; 2004-390329/36.
 KW DR P-PSDB; ADO28954.
 KW PT Novel mammalian G protein coupled receptors, useful for identifying
 KW compounds that modulates diagnosing and treating disease condition
 KW associated with GPCR dysfunction e.g. autoimmune diseases, angina
 KW pectoris, Parkinson's disease.
 KW Claim 13; SEQ ID NO 54; 542bp; English.
 KW The invention relates to human and mouse G protein-coupled receptors

(GPCRs) and nucleic acids encoding them. The invention also relates to sequences at least 90% identical to the GPCR proteins and nucleic acids of the invention; methods of treating, preventing or diagnosing diseases associated with GPCRs of the invention; methods of screening for compounds useful in the treatment of GPCR-related diseases; a transgenic mouse comprising a GPCR gene of the invention; a mouse comprising a mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived from the transgenic mice; kits comprising several GPCR genes; a mutation in a different GPCR gene of the invention; and kits comprising probes which hybridize to GPCR polynucleotides of the invention. The invention further discloses variants of the GPCR polypeptides and vectors comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may be used in the diagnosis, treatment or prevention of a wide variety of diseases including neurological disorders (e.g., Alzheimer's disease, depression, diabetic neuropathy, Parkinson's disease or schizophrenia), disorders of the adrenal gland; disorders of the colon or intestine (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or myocardial infarction); muscular disorders; blood disorders (e.g., anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g., obesity, enzyme deficiency-related diseases or vitamin deficiency-related diseases); and disorders of the kidney, liver, lung, breast, ovary, uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and thyroid (e.g., cancers). The present sequence represents a GPCR-encoding nucleic acid of the invention. Note: The full sequence data for this patent did not form part of the printed specification; those sequences not shown were obtained in electronic format directly from WIP0 at ftp.wipo.int/pub/published_pct_sequences.

Sequence 2273 BP; 503 A; 691 C; 549 G; 530 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,95e-178 Length: 2273
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 12 Gaps: 0

US-09-995-225b-16 (1-353) x ADO28955 (1-2273)

```

QY 1 MetGluHisThrHisAlaHisIleuAlaAlaAsnSerSerLeuSerTrpSerProGly 20
DB 453 ATGGAGCAACAGCAGCCCACTCGCAGCAACAGCTCCCTGCTGTGGTCCGCCGCG 512
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuGlyLeuGly 40
DB 513 TCGGCTCGGGGCTTGGGTTTCGGCCCGTCTCTACTACAGCCCTTGTGTGCTCGGT 572
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnIys 60
DB 573 TTACACGAGCAATATCTTGACAGTGCATCTCTCCAGCTGTGGCAAGAGACAGAG 632
QY 61 SerSerTyrAsnTyrIleuLeuAlaIleAlaAspIleLeuValLeuPhePheIle 80
DB 633 TCTCTTAAACTATCTCTTGGCACTCGTCTGCGACATCTGTCTCTTTTTCATA 692
QY 81 ValPheValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 693 GTGTTGTGCACTTCTCTGTGAGATTTCACTTGAAATGCAATGCAATGCCCTCAGTCCC 752
QY 101 AspLysIleIleGluValLeuGlnPheSerSerIleHisThrSerIleTyrIleThrVal 120
DB 753 GACAGATCAATGAGTCTGGAATTCATCCATCCACACCTCATATGATTAATCTGTA 812
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
DB 813 CGTTAAACATTTGACAGGATATATCATCTGTCCCAACCCGCTCAAGATACACAGCTCTCA 872
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160

```

```

DB 873 TACCAGCCCGCAGCCCGAAGATTTGAAGTTTACATGACCTGTTCTTGACCAAGC 932
QY 161 IleProTyrTyrTrpThrProAsnIleTrpThrGlnAspTyrIleSerThrSerValHis 180
DB 933 ATCCCTTATTAATCTGAGTGGCCCAACATCTGAGCTGAAGACTAATACAGCACTGTGCA 992
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
DB 993 CAGCTCTCATCTGATGCATCCAGCTCAACGCTCACTACCTGAGTCCCTGCTCCATCTTCTTC 1052
QY 201 IleLeuAsnSerIleIleValTyrIleValAlaGlyLysSerAsnPheArgLeuGly 220
DB 1053 ACTTAACTCAATCATTTGTGTGAAGCTCAGAGAGAGCAATTTTGTCTCCGTGCG 1112
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrIle 240
DB 1113 TACTCAGCGGGAAGACCAACCGCATCTTGTTCACCATTACTCCATCTTTGGCACACTT 1172
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
DB 1173 TGGGCCCCCGCATCATGATGATTTCTTACCACTCTATGGGGCGCCATCCAGAACCGC 1232
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
DB 1233 TGGCTGTGACATCATGATTCGACATTCGCAACATGCTAGCCCTTGAAACAGCCATC 1292
QY 281 AsnPheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300
DB 1293 AACTCTTCTCTACTGCTTATCATGACGAGGGTTCGACACAGGACCGCCACACCTC 1352
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 1353 AAGCTTTCTTCAGAGTGCACAGACCACTGTATACAGCAATCATATMACCTTTTCC 1412
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
DB 1413 ATTAACAAGTACCCCTGGATTCGCGGCAAACTCAACATGATCAACAGATGCTGTATC 1472
QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
DB 1473 CAGTATGACAAATGAGAAAACCTATMAAGTATCCCGG 1511

```

RESULT 14
AAS15730
ID AAS15730 standard; cDNA; 1119 BP.
XX
AC AAS15730;
XX
DT 14-FEB-2002 (first entry)
XX
DE DNA encoding chemokine receptor family related protein, NOVI0.
XX
KW NOV, cytostatic, poriasis, nootropic, neuroprotectant;
KW cerebroprotective; hepatic; antiinflammatory; diabetes;
KW haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;
KW differentiation; proliferation; haematopoiesis; wound healing;
KW angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;
KW haemophilia; allergy; Pendred syndrome; skeletal dysplasia;
KW ischaemic injury; neuroepithelial disorder; hepatitis; heart failure;
KW chemokine receptor; chromosome 1; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1119
FT FT
FT FT
FT sig_peptide 1..141
FT FT /tag= a
FT FT /product= "Chemokine receptor related protein, NOVI0"
FT FT /tag= b
FT FT 142..1116
FT FT /tag= c
FT FT /note= "Mature chemokine receptor related protein, NOVI0"

PN W0200170978-A2.
 XX
 XX 27-SEP-2001.
 PF 20-MAR-2001; 2001WO-US009093.
 XX
 XX 20-MAR-2000; 2000US-0190768P.
 PR 20-MAR-2000; 2000US-0190835P.
 PR 22-MAR-2000; 2000US-0190972P.
 PR 22-MAR-2000; 2000US-0191199P.
 PR 24-MAR-2000; 2000US-0191947P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192984P.
 PR 29-MAR-2000; 2000US-0192836P.
 PR 31-MAR-2000; 2000US-0193843P.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX
 PI Taupier RJ, Majumder K, Spederna SK, Smithson G, Mezes PS;
 PI Vermet CAM;
 DR P-PSDB, AAU10067.
 XX
 XX MPI; 2001-639127/73.
 PT Polypeptides and nucleic acids related to chloride channel, insulin-like
 PT growth factor family of proteins, useful for diagnosing and treating
 PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
 XX
 XX Claim 9; Page 42; 151pp; English.
 PS
 XX The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX
 CC polypeptides are useful for treating pathology associated with NOVX
 CC polypeptide, determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX, identifying agents binding to
 CC NOVX and treatment of disorders associated with altered expression of
 CC members of chloride channel-associated proteins e.g. cystic fibrosis and
 CC congenital myotonia. NOVX proteins are useful in treatment of disorders
 CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,
 CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
 CC in one or more organs (e.g. haemophilia, anaemia), Pendered syndrome,
 CC skeletal dysplasia, disorders characterised by altered cell shape,
 CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial
 CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
 CC treatment of disorders of vascular smooth muscle cell differentiation,
 CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are
 CC useful to screen for molecules which inhibit or enhance NOVX activity or
 CC function and are useful as targets for the identifying small molecules,
 CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
 CC differentiation, haemotopoiesis, wound healing and angiogenesis. NOV
 CC sequences are also useful for: identifying a cell or tissue type in a
 CC biological sample; amplifying DNA sequences from very small biological
 CC samples e.g. hair or skin or body fluids and as primers and probes to
 CC identify and/or clone NOVX homologues. NOVX proteins are useful
 CC immunogens to generate antibodies to monitor protein levels and modulate
 CC NOVX activity. Cells comprising the nucleic acids are useful for
 CC producing transgenic animals, for studying the function and/or activity
 CC of NOVX protein and identifying and/or evaluating modulators of NOVX
 CC protein activity. This sequence encodes NOV10 (located on chromosome 1)
 CC related to the chemokine receptor family of proteins, one of 12 NOV
 CC polypeptides described in the method of the invention
 XX
 XX Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;
 SQ

Alignment Scores:

Pred. No.: 2,39e-177 Length: 1119
 Score: 1657.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.20% Indels: 0
 DB: 4 Gaps: 0

US-09-995-225b-16 (1-353) x AAS15730 (1-1119)
 QY 1 MetGluHisThrHisAlaHisIleuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
 DB 1 ATGGAGCACACGACCCACCCACCTCGACGACCAACGCTCGCTGGGGTCCCCGGC 60
 QY 21 SerAlaCyG61YleuGlyPheValProValIvalTyrTyrSerLeuLeuLeuCyGleuGly 40
 DB 61 TCGGCTGGGGCTTGGGGTTTGGTGGCCGGTGTCTTACTACAGCTTGTGCTGCTCGT 120
 QY 41 LeuProAlaAsnIleuLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLeu 60
 DB 121 TTACGACAAATATCTTGAACAGATCATCTCTCCAGCTGGTGGCAAGACAGAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
 DB 181 TCCTCTACAACTATCTTGGGACCTGCGCTGCGGACATCTTGGTCTCTTTTTCATA 240
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 DB 241 GTGTGTGGACTTCTGTGTGGAGATTCATCTTGAACATGACATGCTTCAGTCC 300
 QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 301 GACCAATCATGAAGATGCTGGAAATTCATCATCATCATCATCATCATCATCATCAT 360
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCyHisPProLeuLysTyrHisThrValSer 140
 DB 361 CCGTTAACCATGACGAGTATATCGCTGTGTCACCCGCTCAAGTACACACGAGTCTCA 420
 QY 141 TyrProAlaArgThrArgLysValIleValIleSerValTyrIleThrCySPheLeuThrSer 160
 DB 421 TACCCAGCCCGACCCGGAAGATCATGTAGTATGATTAATCACTGCTTCCTGACACAG 480
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpTrpGluAspTyrIleSerThrSerValHis 180
 DB 481 ATCCCTTATTAATGATGGTGGCCCAACATCGAGAGAAATCATGACGACCTGTGAT 540
 QY 181 HisValLeuIleTrpIleHisCySPheThrValTyrLeuValProCySerIlePhePhe 200
 DB 541 CACGTCCTCATCTGATTCACATGCTTACCGCTTACCTGATGCGCCGCTCCATCTTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
 DB 601 ACTTGAATCATCATCTTGTGTCAAGCTTCAGAGGAAAGACATTTGCTTCGATGG 660
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 661 TACTCACGGGGAAGACCAACCCGATCTTGTACCATTAATCTCTTGTGCACACTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 721 TGGGCCCCCGCATCATCATGATTTCTTTCACACCTTATGGGGCGCCATCCGAAACCC 780
 QY 261 TrpLeuValHisIleMetSerSerIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 781 TGGCTGGTACACATCATGTGCCACATTCGCAACATGTCAGCCCTTGTGAACAGGCATC 840
 QY 281 AsnPhePheLeuTyrCySPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
 DB 841 AACTTCTTCTTCACTGCTTATCATGCAAGCGGTTCCGACATGCAACCGGCAAGCTC 900
 QY 301 LysAlaPhePheLysCyGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 901 AAGGCTTCTTCAAGGCGAGAGCAACCTGTACAGTCTTACACCAATATATACTTTTCC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCyHisMetLeuValTyr 340
 DB 961 ATTAACAAGTAGCCCTGGATCTTCGCGGCAAACTCACTGATCAAGATGCTGGGTAC 1020
 QY 341 GlnTyrAspLysAsnGlyLysProIleLys 350
 DB 1021 CAGTATGACAAATAATGGAATAACCTATTAATA 1050

RESULT 15
ADJ87766
ID ADJ87766 standard; DNA; 1119 BP.
XX AC ADJ87766;
XX 06-MAY-2004 (first entry)
DE G-coupled protein receptor-related protein coding sequence #56.
XX
XX novel protein; G-coupled protein receptor-related protein;
KM cardiomyopathy; atherosclerosis; cell signal processing-related disorder;
KM metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
KM Huntington's disease; epilepsy; anxiety; pain; hypercholesterolemia;
KM obesity; hypertension; Crohn's disease; systemic lupus erythematosus;
KM viral infections; bacterial infection; parasitic infection;
KM hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;
KM Alzheimer's disease; tuberculous sclerosis; hypercalcaemia; cerebral palsy;
KM gene; ds.
XX
XX Unidentified.
OS
PN W02002102321-A2.
XX
XX 27-DEC-2002.
XX
XX 18-JUN-2002; 2002MO-US019522.
PF
XX 18-JUN-2001; 2001US-0298994P.
PR 18-JUN-2001; 2001US-0299134P.
PR 04-OCT-2001; 2001US-00972446.
PR 06-JUN-2002; 2002US-00299134.
PR 07-JUN-2002; 2002US-00298994.
XX
XX (CURA-) CURAGEN CORP.
PA
XX
PI Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;
PI Pena CE, Spaderna SK, Zhong M,
DR WPI; 2003-167441/16.
DR P-PSDB; ADJ87767.
XX
XX
PT New MOX polypeptides and polynucleotides, useful in gene therapy,
PT particularly for treating or preventing e.g. cardiomyopathy,
PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
PT cancer, stroke or pain.
XX
XX
PS Claim 8; SEQ ID NO 201; 378bp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel G-
XX coupled protein receptor-related (MOX) proteins. The DNA and protein
XX sequences of the invention are useful for treating or preventing a MOX-
XX associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
XX associated with cell signal processing and metabolic pathway modulation,
XX or diabetes. The DNA and protein sequences are also useful for the
XX treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
XX pain, hypercholesterolemia, obesity, hypertension, Crohn's disease,
XX systemic lupus erythematosus, viral infections, bacterial infections,
XX parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
XX syndrome, Alzheimer's disease, tuberculous sclerosis, hypercalcaemia, or
XX cerebral palsy. The present DNA sequence encodes a MOX protein of the
XX invention.
SQ Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,398-177 Length: 1119
Score: 1857.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.20% Indels: 0
DB: 10 Gaps: 0

US-09-995-225B-16 (1-353) x ADJ87766 (1-1119)
QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
DB 1 ATGGAGCACACGACGACCACTCCGACGACCACTGCTCTGCGGCCCCGCGC 60
QY 21 SerAlaCy8GlyLeuGlyPheValProValValTyrTyrSerLeuLeuCy8LeuGly 40
DB 61 TCGGCTGCGGCTTGGGTTTGGTGGCCGATGCTTCAAGCTTGTGCTGCTGCTGCT 120
QY 41 LeuProAlaAsnHisLeuThrValHisLeuSerGlnLeuValAlaArgArgGlnlys 60
DB 121 TTACCAAGCAAAATATCTTGACACTGATCATCTCTCCGAGCTGGGCAAGACAGAA 180
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspHisLeuValLeuPhePheHis 80
DB 181 TCCTCTCACTACATATCTTGTGCACTCCGCTGCGACATCTTGCTCTTTTCATTA 240
QY 81 ValPheValAspPheLeuLeuGlnAspPheHisLeuAsnMetGlnMetProGlnValPro 100
DB 241 GTGTTGTGACTTCTCTGTGGAAGATTTCATTTGAACATGACAGATGCTCAGGTCCC 300
QY 101 AspLysHisLeuGlnValLeuGluPheSerSerHisSerSerHisTrpHisVal 120
DB 301 GACAAATCATATGAAAGTCTGAAATCTCATCTCATCCACACTCATATGATTAATGTA 360
QY 121 ProLeuThrHisLeuAspArgTyrHisAlaValCy8HisProLeuLysTyrHisThrValSer 140
DB 361 CGGTTAACCATTTGACAGGTATATGCTGTCTGCGCACCGCTCAAGTACACACGCTTCA 420
QY 141 TyrProAlaAsnThrArgLysValHisLeuSerValTyrHisThrCy8PheLeuThrSer 160
DB 421 TACCAAGCCGACCCGAAAGTCAATGATGATTTTACATCACTGCTCTCTCAACAGC 480
QY 161 IleProTyrTyrTrpTrpProAsnHisLeuThrGlnAspTyrHisSerTrpSerValHis 180
DB 481 ATCCCTATTTACTGTGTGGCCCAATCTGGAAGCTGACATCAACACACTCTGTGCAT 540
QY 181 HisValLeuHisLeuHisValPheThrValTyrLeuValProCysSerHisLeuPhe 200
DB 541 CACGCTCATCTGATCTCACTGCTCACTGCTCACTGCTGCTGCTGCTGCTGCTGCT 600
QY 201 IleLeuAsnSerHisLeuValTyrHisLeuArgHisSerHisPheHisLeuGly 220
DB 601 ACTTGAATCTCAATCATTTGTTACAGCTCAGAGGAGGAAATTTTCTCTCCGTGGC 660
QY 221 TyrSerThrGlyLysThrThrAlaHisLeuPheThrHisSerHisPheHisLeu 240
DB 661 TACTCCAGGGGAAAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 241 TrpAlaProArgHisLeuMetHisLeuTyrHisLeuTyrGlyAlaProHisLeuAsnArg 260
DB 721 TGGGCCCCCGCATCATCATGATCTTTTACCACTTATGGGGCGCCCATCCGAAACCC 780
QY 261 TrpLeuValHisHisMetSerAspHisLeuAsnMetLeuHisLeuAsnThrHisAla 280
DB 781 TGGCTGTACATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 281 AsnPhePheLeuTyrCy8PheHisSerHisPheArgHisThrMetAlaHisAlaThrLeu 300
DB 841 AACCTCTTCTCTACATGCTTCACTCAAGCAAGCGGTTCCGACATGCAAGCCGCAAGCTC 900
QY 301 LysAlaPhePheLysCy8GlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 901 AAGGCTTCTCTCAAGTCCAGAGCAACCTGTACAGTTTCAACCATATTAATCTTTTCC 960
QY 321 IleThrSerSerProTrpHisSerProAlaAsnSerHisValHisMetLeuValTyr 340
DB 961 ATTAACAAGTAAAGCCCTGATCTGCGCGCAACTCACTGATCAAGATGCTGGATGAC 1020
QY 341 GlnTyrAspLysAsnGlyLysProHisLys 350

Db 1021 CAGTATGACAAAATGCGAAAACCTATATAA 1050

Search completed: December 14, 2004, 20:35:21
Job time : 421 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM proteoln - nucleic search, using frame_plus_p2n model

Run on: December 14, 2004, 20:57:16, Search time 81 Seconds
(without alignments)
3097.635 Million cell updates/sec

Title: US-09-995-225b-16
Perfect score: 1872
Sequence: 1 MHTTHAHLAANSLSLWSPG.....CIRKALVQYDKKPKIKVSP 353

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Deiop 6.0, Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Command line parameters:

-MODE=fireme+ p2n.model -DRV=xlh
-Q/cgn2_1/USPTO.spool/US0995225/runat_14122004_110734_18093/app_query.fasta_1.519
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFCG=0
-LOOPEXT=0 -INITS=bits -START=1 -END=-1 -MATRIX=blomsun62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pct -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0995225 @CEN_1_1_54/runat_14122004_110734_18093 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/1/ins/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ins/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ins/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ins/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ins/PTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ins/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241	12.9	1182	1	US-07-629-1041-2
2	241	12.9	1182	1	US-07-629-1041-1
3	238.5	12.7	1194	2	US-08-288-663A-2
4	238.5	12.7	1228	2	US-08-288-663A-3
5	238.5	12.7	1229	4	US-09-016-434-1428
6	232	12.4	1248	4	US-09-545-944-1
7	231.5	12.4	1495	4	US-09-016-434-1190
8	231.5	12.4	1495	4	US-09-023-655-1021
9	231.5	12.4	2156	1	US-08-012-988A-1
10	231.5	12.4	2156	4	US-09-023-655-1247
11	223	11.9	1600	4	US-09-029-027B-1
12	222.5	11.9	2772	4	US-09-919-039-171

13	220	11.8	2706	3	US-08-676-351-1	Sequence 1, App1
14	219.5	11.7	1217	4	US-09-016-434-1446	Sequence 1446, Ap
15	219	11.7	1452	1	US-08-149-092A-3	Sequence 3, App1
16	219	11.7	1452	1	US-08-911-245-3	Sequence 3, App1
17	219	11.7	1452	1	US-08-553-058C-3	Sequence 3, App1
18	219	11.7	1452	2	US-08-514-451A-3	Sequence 3, App1
19	219	11.7	1452	3	US-09-170-331-3	Sequence 3, App1
20	219	11.7	1452	3	US-09-510-473-3	Sequence 3, App1
21	219	11.7	1452	4	US-09-048-916B-3	Sequence 3, App1
22	218	11.6	1137	3	US-09-045-583-6	Sequence 6, App1
23	218	11.6	1137	3	US-09-534-185-6	Sequence 6, App1
24	218	11.6	1664	4	US-09-045-583-4	Sequence 4, App1
25	218	11.6	1664	4	US-09-534-185-4	Sequence 4, App1
26	217	11.6	1330	3	US-08-147-592A-5	Sequence 5, App1
27	217	11.6	1330	3	US-08-292-694A-5	Sequence 5, App1
28	217	11.6	1567	3	US-08-889-108-16	Sequence 16, App1
29	217	11.6	1567	5	PCT-US94-1035B-16	Sequence 16, App1
30	217	11.6	2600	4	US-08-986-209A-1	Sequence 1, App1
31	217	11.6	2706	2	US-08-454-549-1	Sequence 1, App1
32	217	11.6	2706	3	US-08-454-552-1	Sequence 1, App1
33	216.5	11.6	1225	4	US-09-016-434-1235	Sequence 1235, Ap
34	214.5	11.5	1317	1	US-08-153-848-45	Sequence 45, App1
35	214.5	11.5	1317	1	US-09-299-843A-45	Sequence 45, App1
36	214.5	11.5	1317	3	US-09-088-337B-45	Sequence 45, App1
37	214.5	11.5	1317	5	PCT-US93-11153-45	Sequence 45, App1
38	213.5	11.4	1275	4	US-09-341-446B-7	Sequence 7, App1
39	213	11.4	1664	4	US-09-582-224A-5	Sequence 5, App1
40	213	11.4	1664	4	US-09-023-655-1213	Sequence 1213, Ap
41	213	11.4	1679	3	US-09-517-605-14	Sequence 14, App1
42	213	11.4	1737	1	US-08-202-056-4	Sequence 4, App1
43	213	11.4	1737	1	US-08-076-093A-3	Sequence 3, App1
44	213	11.4	1737	1	US-08-701-265-3	Sequence 3, App1
45	213	11.4	1737	2	US-08-284-586-3	Sequence 3, App1

ALIGNMENTS

RESULT 1
US-07-629-1041-2
Sequence 2, Application US/076291041
Patent No. 5288621
GENERAL INFORMATION:
APPLICANT: Gershangorn, Marvin C
TITLE OF INVENTION: PITUITARY TRH RECEPTOR
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/629,1041
FILING DATE: 19901214
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: CRF D - 995
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 2 :
SEQUENCE CHARACTERISTICS:
LENGTH: 1182 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (cDNA)
US-07-629-1041-2

Alignment Scores:

Pred. No.:	9,88e-17	Length:	1182
Score:	241.00	Matches:	85
Percent Similarity:	41.78%	Conservative:	70
Best Local Similarity:	22.91%	Mismatches:	121
Query Match:	12.87%	Indels:	96
DB:	1	Gaps:	15

US-09-995-225b-16 (1-353) x US-07-629-1041-2 (1-1182)

```

QY 19 ProGlySerAlaCysGlyLeuGlyPhe---ValProValValTyrTyrSerLeuLeu 37
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 46 CCACACGACGCTGTGGCCCTCGAGTACGAGTGGTGTACCTTACTTGTGTGCTATTAT 105

QY 38 Cys---LeuGlyLeuProAlaAsnIleLeuThrValIleIleuSerGlnLeuValAla 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 106 TGTGAGCTGGGCACTTGTGGCCACATCATGTAGTCTGTGGTGTGATGAGAACAAAC 165

QY 57 ArgArgGlnIlySerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuVal 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 ATGAGAACCCCTCAACACTGT---TACCTGTAAAGTCTGTGGTGGCAGATCTCATGGTT 222

QY 77 LeuPhePheIleValPheValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMet 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 CTG-----GTGGCTGCAGAGACTC 240

QY 97 ProGlnValProAspIlyIle-----GTGGCTGCAGAGACTC 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 CCCAACATTAACCCAGATATCTATGGTCTGTGGTCTATGGCTATGTGGCTGCTCTG 300

QY 104 ILeGluValLeuGlnIlyPheSerIleIleThrSerIleTyrIleThrValProLeuThr 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 ATTACATATCTCCAGTACCTAGGCAATTATGATCTTCAATGTTCAATTAACGGCTTTTACC 360

QY 124 ILeAspArgTyrIleAlaValCysHisProLeuIlySerIleThrValSerTyrProAla 143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 ATGGAAGGATACATACATCTGTCACTCCCATTAAGCCAGTTTCTGCACAGTTTCTCC 420

QY 144 ArgThrArgIlyValIleValSerValTyrIleThrCysPheLeuThrSerIleProTyr 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 AGAGCCAAAATAATCATCTTGTCTGTGGCC-----TTCACATCCATTATACGT 471

QY 164 TyrTyrTyr-ProAsnIleTyrThrGlnAspTyrIleSerThrSerVal---HisHisVal 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 ATGCTGTGTTCTCTCTGCTG-----ATCTCAATCATGACCACTTACAAAAC 519

QY 182 IleuIleTyrIleHisCys-----PheThrValTyrIle 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 GCTGTGTGTGTTTCC-TGTGGCTACAAAGATCTCCAGAACTACTACTACCACTTATTA 578

QY 193 uValProCysSerIlePhePheIleLeuAsnSerIle-----IleValTyr----- 208
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 579 AATGACCTTGGTGTCTTTATGTGTGCAATGATCTGTGGCACTGTGCTTATGATT 638

QY 208 ----- 208

Db 639 TATAGCTAGAACTCTCTTCTTAACCCCATTCCTCAGACCCCTAAAGAAACTTAAGAT 698

QY 209 -----IlySerArgIlySerArgPheArgIly----- 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 699 GTGGAATAATGATCATTCATGACAAAGAAATTGATTTAATGACCAACAAAGATG 758

QY 221 -----TyrSerThrGlyIlySerThrThrAlaIleuPheThrIleThrse 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 759 CTTCAACAGCACTGATCTTCCAGAAAGAGCTCAACAAAGCTCGCAGTGTGTAT 818

QY 235 rIlePheAlaThrIleuThrAlaPro---ArgIleIleuThrIleuThrIleuThrIleuThr 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 819 TCTGTTGGCTTTTATGATGCCCTACAGACTCTTGTGTGTCACTATTCTCTC 878
QY 254 yAlaProIleGlnAsnArgTyrLeuValHisIleMetSerAspIleAlaAsnMetLeuAl 274
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 879 CAGCCCTTCCAGAAAATTGTCTCTGCTC-----TTTGGCAATTTGCAT 926

QY 274 aleuLeuAsnThrAlaIleAsnPheLeuTyrCysPheIleSerIlySerArgPheArgTh 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 927 TTATCTCAACAGTGCATCAACCGATATTACAACTCATGTCTCAAAAATTTCGT-- 984

QY 294 rMetAlaAlaThrIleuIlyValPhePheIlyCysGlnIlyGlnProValGlnPheThy 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 985 -----GGAGCCTTCCAGAACTGTGCATTTGCAAGAGAACCCACAGAAAAGC 1034

QY 314 rThrAsnHisAsnPheSerIleThrSerSer 324
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1035 TGTAACTACAGTGTGGCCCTAAATTAACGT 1065

```

RESULT 2

```

US-07-629-1041-1
Sequence 1, Application US/076291041
Patent No. 5288621
GENERAL INFORMATION:
APPLICANT: Gershengorn, Marvin C
APPLICANT: Straub, Richard E
TITLE OF INVENTION: PITUITARY TRH RECEPTOR
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/629,1041
FILING DATE: 19901214
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: CRF D - 995
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 268-1951
TELEFAX: (203) 268-1951
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1752 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (cDNA)
US-07-629-1041-1

```

Alignment Scores:

Pred. No.:	1.79e-16	Length:	1752
Score:	241.00	Matches:	85
Percent Similarity:	41.78%	Conservative:	70
Best Local Similarity:	22.91%	Mismatches:	121
Query Match:	12.87%	Indels:	96
DB:	1	Gaps:	15

US-09-995-225b-16 (1-353) x US-07-629-1041-1 (1-1752)

```

QY 19 ProGlySerAlaCysGlyLeuGlyPhe---ValProValValTyrTyrSerLeuLeu 37
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 CCACACGACGCTGTGGCCCTCGAGTACGAGTGGTGTACCTTACTTGTGTGCTATTAT 363

```

```

QY 38 Cys---LenglyleuProAlaAsnIleuThrValIleIleuSerGlnLeuValAla 56
Db 364 TGTGACCTGGGCACTGTGGCAATCATGTAGTCTGGTGCATGAGAACAAAGCAC 423
QY 57 ArgArgGlnIleuSerSerIleuThrValIleIleuSerGlnLeuValAla 76
Db 424 ATGAGAAACCCCTCAAACTCT---TACCTGTAAAGTCTGGTGCAGAGTCTCATGGTT 480
QY 77 LeuPhePheIleValPheValAspPheLeuGlnAspPheIleuAsnMetGlnMet 96
Db 481 CTG-----CTG-----GTGCTGCAGAGCACTC 498
QY 97 ProGlnValProAspIleuSerIleuThrValIleIleuSerGlnLeuValAla 103
Db 499 CCCAACAATACCGACGATCATGTGCTGGTCTGATGCTGATGCTGCTGCTGCTGCTG 558
QY 104 IleGlnValIleuGlnPheSerSerIleuThrValIleIleuSerGlnLeuValAla 123
Db 559 ATTACATATCTCCAGTACCTAGGCAATTAATGATCTTCAATGATCAATGAGGCTTTACC 618
QY 124 IleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSerTyrProAla 143
Db 619 ATGAAGAAGTACATACATCTGTCAACCCCATCAAGCCAGTTCTCTGACGTTTCTCC 678
QY 144 ArgThrArgIleValIleValSerValTyrIleThrCysPheLeuThrSerIleProTyr 163
Db 679 AGAGCCAAATAATCATCATCTTGTCTGGGCC-----TTCAATCATTTACTGCT 729
QY 164 TyrTrpTyrProAsnIleThrPheGlnAspTyrIleSerThrSerVal---HisHisVal 182
Db 730 ATGCTCGTGTCTTCCGCTG-----ATCTCAATCATGACGACTCAAAAC 777
QY 182 IleuIleTrpIleHisCys-----PheThrValTyrIle 193
Db 778 GCTGTGTGGTTC---TGCGCTACAAGATCTCCAGAACTACTACTCACCCTATTAACT 836
QY 193 ValAlProCysSerIlePhePheIleLeuAsnSerIle-----IleValTyr----- 208
Db 837 AATGACCTTGGTGTCTTATGTGTGCAATGATCTGCGCACTGTGCTTATGAGATT 896
QY 208 ----- 208
Db 897 TATAGCTAGAATCCTCTTCTTAACCCCATCTTCAGACCTTAAGAAACTCTAAGAT 956
QY 209 -----LysLeuArgArgIleSerAsnPheArgLeuArgGly----- 220
Db 957 GTGGAATAATGACTCAATCATCAAGAAATTTGAATTTAATGCCACCAAGATG 1016
QY 221 -----TyrSerThrGlySerThrAlaIleLeuPheThrIleThrIse 235
Db 1017 CTTCAACAGACGTATCTTCTTCAAGAAAGTCAAGATGCTGCGAGTGTGTAAT 1076
QY 235 rIlePheAlaThrLeuTyrAlaPro---ArgIleIleMetIleuTyrHisLeuTyrGln 254
Db 1077 TCTGTTGCCCTTTATGATGATGCCCTACAGACACTTGTGTGCTCAACTTCTCTC 1136
QY 254 ValAlaProIleGlnAsnArgTyrLeuValHisIleMetSerAspIleAlaAsnMetLeuAl 274
Db 1137 CAGCCCTTTCAGAAATATGTTCTTCTC-----TTTTCAGAAATTTTCAT 1184
QY 274 AluLeuAsnThrAlaIleAsnPhePheLeuTyrCysPheIleSerIleValArgPheArgTh 294
Db 1185 TTATCTCAACAGGCTCAACCAAGATTTACAACTCATGCTCAAGAAATTTCT- 1242
QY 294 rMetAlaAlaAlaThrLeuValPhePheLysCysGlnLysGlnProValGlnPheTyr 314
Db 1243 -----GAGAGCTTCAGAAAGCTTCGCAATTCGACAGAGCAAGACCCACAGAAAGC 1292
QY 314 rThrAsnHisAsnPheSerIleThrSerSer 324
Db 1293 TGCTAACTACAGTGTGGCCCTTAATTAACAGT 1323

```

```

US-08-288-663A-2
; Sequence 2, Application US/08288663A
; Patent No. 5879896
; GENERAL INFORMATION:
; APPLICANT: HINDU, Shuji
; APPLICANT: HOSODA, Masaki
; APPLICANT: ONDA, Haruo
; TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSER: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,663A
; FILING DATE: 09-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 198309/1993
; FILING DATE: 10-AUG-1993
; APPLICATION NUMBER: 286986/1993
; FILING DATE: 16-NOV-1993
; APPLICATION NUMBER: 325215/1993
; FILING DATE: 22-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 44612
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-288-663A-2
; Alignment Scores:
; Pred. No.: 1,88e-16 Length: 1194
; Score: 238.50 Matches: 85
; Percent Similarity: 41.21 Conservative: 72
; Best Local Similarity: 22.31 Mismatches: 130
; Query Match: 12.74 Indels: 95
; DB: 2 Gaps: 13
US-09-995-225b-16 (1-353) x US-08-288-663A-2 (1-1194)
QY 2 GlnHisThrHisAlaHisLeuAlaAlaAsnSerSerIleuSerTyrTrpSerProGlySer 21
Db 32 AAACACAGGCTTCAGCCAC-----GAGCAGTGTGGCTTGAATACC 73
QY 22 AlaCysGlyLeuGlyPheValProValTyrTyrSerLeuLeuLeuValAlaArgArgGlnLysSer 41
Db 74 AGS-----TGTCACCATCTTACTTACTTATTTGTGC-CTGGGCATT 120
QY 42 ProAlaAsnIleuThrValIleIleuSerGlnLeuValAlaArgArgGlnLysSer 61
Db 121 GTAGCAACATCTTGTAGTCTGTGCTGATGAGAACCAAGACATGAGACCCCA 180
QY 62 SerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleuValLeuPhePheIleVal 81

```


[illegible]

```

ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1428:
SEQUENCE CHARACTERISTICS:
LENGTH: 1229 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 9577631
US-09-016-434-1428

Alignment Scores:
Pred. NO.: 1,976-16 Length: 1229
Score: 238.50 Matches: 85
Percent Similarity: 41.21 Conservative: 72
Best local Similarity: 22.31 Mismatches: 130
Query Match: 12.74 Indels: 95
DB: 4 Gaps: 13

US-09-995-225B-16 (1-353) x US-09-016-434-1428 (1-1229)
QY 2 GlnHstHthtAlahtIsleuAlaAaenSerleuSerTyrtpSerProglySer 21
   ::::: :::::
Db 44 AAACAAGCTTCAGCCAC-----GAGCAGTGTGCTTGAATAC 85
QY 22 AlAcGgLyLeuGlYpHeValProValValTyrTyrSerleuLeuCyLeuGlYleu 41
   ::::: :::::
Db 86 AGG-----TGTTCACCACTTACTTGTCTAATTTTGTGGC-CTGGCATT 133
QY 42 ProAlaSnlleuThrValIleIleuSerGlnleuValAlaArgArgGlnlySer 61
   ||||| ::::: :::::
Db 133 GTAAGCAACTATGTGTAGTCTCGGTTCATGAGAACCAACACATAGACCCACCA 192
QY 62 SerTyrAsnTyrlleuAlaIleuAlaAlaIleuValIleuValIleuPheIleVal 81
   ::::: ::::: :::::
Db 193 AACTGCTC-TACTCTGTGAGCTCGGACATGATCTCATCGGTCTTG----- 233
QY 82 PheValAspPheleuGlYuaPheIleuAenMetGlnetProGlnValProAsp 108
   ::::: :::::
Db 238 -----GTGCCCGCAGGCTCCCAACATACAGAC 266
QY 102 lyeIle-----IleGlnValleuGln 108
   |||||
Db 268 AGATATCTACGAGTTCTCGGCTCATGAGCATATGTGGATGCTCTGATTACTTACCTCAG 322
QY 109 PheSerIleHstHstSerIleTrpIleThrValProleuThrIleAspArgTyrIle 128
   ||||| ::::: |||||

```

```

Db      328 TATTGGGAATTATGATCCTTTGTCATATACAGCCTTACCATTTAGAGGTACATA 387
QY      129 AlAlValCySHsProluYrYrHisThValSerYrProAlaArgHisVal 148
      388 GCAATCTGTACCCCAAGGCCAGTTTCTGCACATTTTCCAGAGCCAAAGAT 447
QY      149 lIeValSerValYrIleThrCysPheLeuThrSerIleProYrYrTrp 166
      448 ATCATCTTGTCTGGGCT-----TTCAATCTTTTACTGATGCTCTGTTCTTC 498
Db      167 -----ProAnIleTrpThr--GluAspYrIleSerThrSerValHisVal 182
QY      499 TTGCTGATCTCATATATTAGCACCACAAAGATGCTATGTGATCTCTGGCTACAG 558
Db      183 LeuIleTrpIleHisCysPheThrValYrLeuValProCysSerIlePhePheIleu 202
      559 ATCTCCAGGAATTACTACTACCTATTACTTAATGACTTTGGTGTCTTATGTGTG 618
QY      203 AsnSerIleIleValYrYrIleu----- 210
      619 CCAATGATCTGGCTACCGTCTCTATGATTCATAGCTAGATCCTTTTCTTAATCCC 678
QY      211 ----- 213
      679 ATTCCTTCAGATCTTAAGAAACTTAAGACATGGAATAATGATTCACCCATCAGAAC 738
Db      214 SerAsn-----PheArgLeuArgIleYrSerThrIle 224
      739 ACAAACTGATTAATTAATACCTTAATAGATGTTTCAACAGCACAGATATCTTCAAGAA 798
QY      225 LysThrThAlaIleLeuPheThrIleThrSerIlePheAlaThrLeuTrpAlaPro 243
      799 CAGGTCCACCAAGATGCTGGAGGTGTGTAATCTGTGGCTTTTGAATGCCCTTAC 858
Db      244 ArgIleIleMetCysLeuYrHisLeuYrGlyAlaProIleGluAsnArgTrpLeuVal 263
      859 AGACCTCTAGGTGTGTCAACTCATTTCTCTCCAGCTTTCACAAATAATGTGTTTGG 918
QY      264 HisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIleAsnPhePhe 283
      919 CTC-----TTTTCAGAAATTGTCATTTATCTAACAAGTCCATCAACCCGGTG 966
Db      284 LeuYrCysPheIleSerYsArgPheArgThrMetAlaAlaAlaThrLeuYsAlaPhe 303
      967 ATTTCAAATCTCATGTCCCAAGAAATTCGCT-----GCAGCCTTCAGAAAGCTC 1014
QY      304 PheYsCysGlnIleGlnProValGlnPheYrThrAsnHisAsnPheSerIleThrSer 323
      1015 TGCACCTGACAGCAGAAAGCCAAAGAAACCTGCTAATCAAGTGTGGCCTTAATTAAC 1074
QY      324 Ser 324
      1075 AGC 1077
Db

```

```

; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1248
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-545-944-1

Alignment Scores:
Pred. No.: 1,04e-15 Length: 1248
Score: 232.00 Matches: 79
Percent Similarity: 44.32% Conservative: 77
Best Local Similarity: 22.44% Mismatches: 112
Query Match: 12.39% Indels: 84
DB: 4 Gaps: 15

US-09-995-225B-16 (1-353) x US-09-545-944-1 (1-1248)

QY      27 PheValProVal---ValYrYrSerLeuLeuCysIleuGlyLeuProAlaAsnIle 45
      133 TTCCTCCCGGTCTGTGTGTATGTGCAATTTTGTGTGGGGGTCAATTGGCAATGTC 192
Db      46 LeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnYsSerSerYrAsnYr 65
      193 CTGGTGGCTGGTGTATCTGTGACGACACAGGTATGAAGAGCCCACTAC---TAC 249
QY      66 LeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIleValPheValAspPhe 85
      250 CTCCTACCTGGCGGTCTGTGACCTCTGTGGTCTG-----CTCCTTGAATG 297
QY      86 LeuLeuGluAspPheIleLeuAsnMetGlnMetPro----- 97
      298 CCCCTGGAGGTCTAGATGTGGCGCACTACCTTTCTTGGGCCCGGTGGCTGC 357
Db      98 GlnValProAspYsIleIleGlnValLeuGlnPheSerSerIleHisThrSerIleTrp 117
      358 TACTTCAAGACGGCCCTCTTGAAGACGGTGTGGCTTCCGCTCATC-----CTAGC 408
QY      118 IleThrValProLeuThrIleAspArgYrIleAlaValCysHisProluYrYrHis 137
      409 ATCAACACC--GTACAGGTGAAGCGCTACGTGCATCTCAACCCGTTCCGGCCAAA 465
Db      138 ThrValSerYrProAlaArgThrArgIleValIleValSerValYrIleThCysPhe 157
      466 CTGAGAGCACCGCGCGCGCGCCCTCAGATCTCGGCATCGTGGGCTTCTCCGTG 525
QY      158 LeuThrSerIlePro-----TyrYrTrpTrpProAsn--- 168
      526 CTCCTTCCCTGCCCAACACAGCATCCATGATCAAGTTCCACTTCCCAATGGG 585
Db      169 -----IleTrpThrGluAspYr 174
      586 TCCCTGTGCCAGGTGGGCCACTGACGTCATCAAGCCCATGTGG----- 633
QY      175 lIeSerThrSerValHisValLeuIleTrpIleHisCysPheThrValYrLeuVal 194
      634 -----ATCTCAATTTTCATCATCCAGGTACCTCTTCCATTTACTCTCTC 681
Db      195 ProCysSerIlePhePheIleLeuAsnSerIleIleValYrYsLeuArg- 212
      682 CCATGACTGTCACAGTGTCTCTACTACTGACCTGACACTAAGAAAGAACAA 741
QY      213 -----LysSerAsnPheArgLeuArgIleYrSerThrIleGlySerThr 227
      742 TCTCTTAGCAATGAAGGAATGCAATATTCAAAGCCCTGCAGAAATATAGTCAAC 801
QY      228 AlAlleuPheThrIleThrSerIlePheAlaThrLeuTrpAlaProArgIleIleMet 247
      802 AAGATGCTGTGTGTGTGCTTGAAGTGTGATCTGTGGGCCCGGTCCACATTTGAC 861
QY      248 lIleuYrHisLeuYrGlyAlaProIleGlnAsnArgTrp----- 261
      862 CCACTCTTCTTCACTTT-----GTGAGAGAGTGAAGTATCCCTGGCTGCT 909

```

QY 262 -----leuValHsileMetSerApplleAlaAsmMetleuAlaLeuAsnThr 278
DB 910 GGTTCACCTCTCTCAGTGTGTCAGTGC-----TTATTCTACCTGAGCTCA 960
QY 279 AlaileasnphepheleuTyrcyspheileSerlysarpheargThrmetAlaAla 298
DB 961 GCTGTCAACCCCATTTACTTAACTTACTCTGCGCGCTTCCAGGACAGATCCAGAA 1020
QY 299 ThrleuValAlaPhepheleTyrcysGlnLysGlnProValGlnPheTyThrAsnHis 318
DB 1021 GTGATCTCTCTTCCACAAA-----CAGTGGCACTCCAGCATGAC 1062
QY 319 PheSerleTherSerProTrrPileserProAla 330
DB 1063 -----CCACAGTGGCACCTGCC 1080

RESULT 7
US-09-016-434-1190
Sequence 1190, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1190:
SEQUENCE CHARACTERISTICS:
LENGTH: 1495 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g179984
US-09-016-434-1190

Alignment Scores:
Pred. No.: 1,55e-15 Length: 1495
Score: 231.50 Matches: 81
Percent Similarity: 44.81% Conservative: 57
Best Local Similarity: 26.30% Mismatches: 105
Query Match: 12.37% Indels: 65
DB: 4 Gaps: 13

US-09-995-225B-16 (1-353) x US-09-016-434-1190 (1-1495)

QY 22 AlaCysGlyLeuClYpheValProValValTyTySerleuLeuLeuCysleuGlyLeu 41
DB 91 GCTTTGGGGCCCAACCTGCTCCCTCG-----TACTCTCTGGTATTGTCACTGGCCG 147
QY 42 ProAlaAsnIleleuThyValIleileSerGlnLeuValAlaArgGlnLysSer 61
DB 148 GTTGGAAACATCTGGTGTCTCTGCTTGTGCAATACAAAGGCTAAAAAAAGATGAC 207
QY 62 SerTyAsnTyThrleuValAlaAlaAspIleleuValleuValPhePheleVal 81
DB 208 AGCATC---TACTCTGAACCTGGCCATTTCTGACCTGCTTCTGTTACAGCTTCC 264
QY 82 Phe---ValAspHe---leuLeuGluAAspPheIleleuAsmMetLmetProGlnVal 99
DB 265 TTCTGATGATGACTCAAGATTGAAGATGACTGGGTTTGTGATGCCATG----- 315
QY 100 ProAspPheIleleuGluValleuGluPheSerSerleIleThSerleIleThr 119
DB 316 ---GTAAAGTCTCTCTGCGTTTATTATACACAGGCTTACAGCAGATCTTTTCAATC 372
QY 120 ValProleuThrIleAspArgTyriIleAlaValCysHis-----ProleuTyTy 136
DB 373 ATCTGCTGACGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 432
QY 137 HisThValSerTyProAlaArgThrArgValIleValSerValTyriIleThrCys 156
DB 433 CGGACCGCATCTTTGTGTCATCACACCATCATCATTTGGGCTCGGCATC----- 486
QY 157 PheleuThSerlePro---TyTyTyTrpProAsnIleThrGluAspTyriIle 175
DB 487 ---TTGGCTTCCATGCGAGGCTTATCTTTCCAAACCATGGAATTCATCAACAC 543
QY 176 SerThSerValHisHis-----ValleuIleTrpIleHisCysPheThVal 191
DB 544 ACCTGACGCTTCACTTCTCAGAAAGCTACGAGATGGAAGCTGTTCAAGGCTG 603
QY 182 -----ValleuIleTrpIleHisCysPheThVal 191
DB 604 AAACGAACTCTTGGGCTGTATGCTTTGTGTGATGATGATGATGATGATGATGATG 660
QY 192 TyrleuValProCysSerleIlePhePheIleleuAsnSerleIleValTyTyLysLeuArg 211
DB 661 -----GGATTTTAAAGATTCTGCTTAA 684
QY 212 ArgLysSerAsnPharArgLeuArgGlyTySerThrGlyLysThrAlaIleleuPhe 231
DB 685 CGACCAATGAG-----AAGAAATCCAAAGCTGTCGTTGATTTT 726
QY 232 ThrIleThSerlePheAlaThrleuThrAlaProArgIleIleMetIleleuTyHis 251
DB 727 GTCATCATGATCATCTTTTCTTTTCTTTGGACCCCTTACAACTTAACTTATTTCT 786
QY 252 leuTyGlyAlaProle-----GlnAsnArgTrpLeuValHisIle 265
DB 787 GTTTCAAAGACTTCTGTTCACCCATGAGTGTGACAGAGCAGCATTTG---GACCTG 843
QY 266 MetSerAspIleAlaAsmMetleuAlaLeuLeuAsnThrAlaIleAsnPharPheleuTy 285
DB 844 GCTGTGCAAGTACGAGAGTATCGCTTACAGCACTGCTGTGTAACCAAGATGATAC 903
QY 286 CysPheIleSerTyAspPheArg 293
DB 904 GCCTTCGTGTGAGAGGTTCCGG 927

RESULT 8
US-09-023-655-1021
Sequence 1021, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Sellhamer

Db	373	ATCTGCTGACGATTACAGATGACCTGGCCATCGTCCAGCCGCTGTTGCTTGGCGGCA	433
Qy	137	HisThValSerTyrrProAlaargThnArgLysValIleValSerValTyrrIleThnCys	156
Db	433	CGGACCGTCATCTTTGGTGTGCATCAGCAGCATCATTTGGAGCCGCGCATC-----	486
Qy	157	PhelLeuThrSerIlePro---TyrrTyrrTlPrProAsnIleThrPhnArgLysApTyrrIle	175
Db	487	----TGGCTTCATGCGCAGGCTTATCTTTCCAAAGACCCAAATGGAAATTCCTACTCAC	543
Qy	176	SerThrSerValHisHis-----	181
Db	544	ACCTGCAGCCTTCATCTTCTCCACGAAAGCTACGAGAGTGAAGCTGTTTCAGGCTCG	603
Qy	182	-----ValleuIleTrrPleHisCysPheThnVal	191
Db	604	AAACTGACACTTTTGGCGTGGTATATGCTTGTGGTATGATCATTCGTACACACAC	660
Qy	192	TyrrLeuValProCysSerIlePhePheIleuAsnSerIleIleValTyrrLysLeuArg	211
Db	661	-----GGATTTATTAAGTTCTGCTAA	684
Qy	212	ArgLysSerAsnPheArgLeuArgLysTyrrSerThnArgLysThnThrAlaIleLeuPhe	231
Db	685	CGACCAAAATGAG-----AAGAAATCCAAAGCTGCGCTTGATTT	726
Qy	232	ThrIleThnSerIlePheAlaThnLeuThrLeuThrAlaProArgIleIleMetCileuTyrrHis	251
Db	727	GTCATCATGATCATCTTTTCTCTTTTGGACCCCTTCATTTGACTATGATCTTATTTCT	786
Qy	252	LeuTyrrGlyAlaProIle-----GlnAsnArgTyrrLeuValHisIle	265
Db	787	GTITTCACAGCTTCTCTGTTCACCCATGATGATGTCAGACAGACAGACATTTG---GACCTG	843
Qy	266	MetSerArgPleIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPheLeuTyrr	285
Db	844	GCTGCGACAGTACGAGAGTATGCTTACACGACGCTGCTGTTCATACCACTGATCTAC	903
Qy	286	CysPheIleSerLysArgPheArg	293
Db	904	GCCTTCGTTGGTGAGAGGTTCCGG	927
RESULT 9			
US-08-012-988A-1			
: Sequence 1, Application US/08012988A			
: Patent No. 5652133			
: GENERAL INFORMATION:			
: APPLICANT: Murphy, Philip M.			
: TITLE OF INVENTION: Cloning and Expression of Human			
: TITLE OF INVENTION: Macrophage Inflammatory Protein-1 alpha (MIP-1			
: NUMBER OF SEQUENCES: 2			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: Townsend and Townsend Kourile and Crew			
: STREET: One Market Plaza, Stewart Tower, Suite 2000			
: CITY: San Francisco			
: STATE: California			
: COUNTRY: USA			
: ZIP: 94610			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Floppy disk			
: COMPUTER: IBM PC compatible			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: PatentIn Release #1.0, Version #1.25			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/012,988A			
: FILING DATE: 19930128			
: CLASSIFICATION: 435			
: ATTORNEY/AGENT INFORMATION:			
: NAME: Weber, Kenneth A.			
: REGISTRATION NUMBER: 31,677			
: REFERENCE/DOCKET NUMBER: 15280-118			
: TELECOMMUNICATION INFORMATION:			


```

TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2156 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: primer bind
LOCATION: 259..275
FEATURE:
NAME/KEY: primer_bind
LOCATION: 259..275
FEATURE:
NAME/KEY: complement (868..884)
LOCATION: complement (868..884)
FEATURE:
NAME/KEY: CDS
LOCATION: 63..1128
US-08-012-988A-1
Alignment Scores:
Pred. No.: 2,69e-15 Length: 2156
Score: 231.50 Matches: 81
Percent Similarity: 44.81% Conservative: 57
Best Local Similarity: 26.30% Mismatches: 105
Query Match: 12.37% Indels: 65
DB: 1 Gaps: 13
US-09-995-225B-16 (1-353) x US-08-012-988A-1 (1-2156)
QY 22 AlaCysGlyLeuValProValValTyrTyrSerLeuLeuLeuCysLeuGlyLeu 41
DB 153 GCCTTGGGGCCCACTGCTGCCCCCTCTG--TACTCTGTGATTGTCATTGGCCTG 209
QY 42 ProAlaAsnIleuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnSer 61
DB 210 GTTGGAAACATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 269
QY 62 SerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPheIleVal 81
DB 270 AGCATC--TACTCTCTGAACCTGGCCATTCTGACCTGCTCTCTCTGACGCTTCCC 326
QY 82 Phe--ValAspPhe--LeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnVal 99
DB 327 TTTGTGATGACTACAGATTGAAGATGACTGGGTTTGGTATGCCATG-- 377
QY 100 ProAspIleIleGlnValLeuGlnPheSerSerIleIleAsnSerIleTyrIleThr 119
DB 378 ---TGTAAGTCTCTCTGCGTTTATTACACAGGCTTGACAGCGAGTCTTTTCATC 434
QY 120 ValProLeuThrIleAspArgTyrIleAlaValCysHis-----ProLeuLysTyr 136
DB 435 ATCTCTCTGACGATGACAGGTACCTGGCCATGCTGCACGCCGTTGGCTTGGCGGCA 494
QY 137 HisThrValSerTyrProAlaArgTyrArgValIleValSerValTyrIleThrCys 156
DB 495 CGAGCGGTCACTTTGGTGTGCATCAGCAGCATCATTTGGGCTTGGCGCATC----- 548
QY 157 PheLeuThrSerIlePro--TyrTyrTyrProAsnIleTyrPheGlnAspTyrIle 175
DB 549 ---TTGGCTTCCATGCGCAGGCTTATCTTTCCAAAGCCAAATGGGAATTCATCACCAC 605
QY 176 SerThrSerValHisHis-----ValLeuIleTyrIleHisCysPheThrVal 181
DB 606 ACCTGACGCTTCACCTTCTCTCAGAAAGCTACGAGTGAAGCTGTTTCAGGCTCTG 665
QY 182 -----ValLeuIleTyrIleHisCysPheThrVal 191
DB 666 AAACGTAACCTCTTGGCGCTGATGCTTGTGGTCAAGATCAGTGTCTACACA--- 722
QY 192 TyrLeuValProCysSerIlePhePheIleLeuAsnSerIleIleValTyrLysLeuArg 211
DB 723 -----CGGATTATAAAGATTCTGCTTAAGA 746

```

```

QY 212 ArgLysSerAsnPheArgLeuArgLysTyrSerThrGlyLeuThrAlaIleLeuPhe 231
DB 747 CGACCAAAATGAG-----AAGAAATCCAAAGCTGCTCGTTGATTTT 788
QY 232 ThrIleThrSerIlePheAlaThrLeuTyrAlaProArgIleMetIleLeuTyrHis 251
DB 789 GTCATCATGATCATCTTTTCTCTCTTTTGACCCCTTACATTTGACTTACTTATTCT 848
QY 252 LeuTyrGlyAlaProIle-----GlnAsnArgTyrLeuValHisIle 265
DB 849 GTTTCCAAGACTCTCTGTTCAACCATGAGTGAGAGACAGACATTTG--GACCTG 905
QY 266 MetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPheLeuTyr 285
DB 906 GCTGTGCAAGTACGAGGATGATCGCTACACGCACTGCTGTGTCAACCATGATCTAC 965
QY 286 CysPheIleSerLysArgPheArg 293
DB 966 GCCTTCGTGTGAGAGGTTCCGG 989
RESULT 10
US-09-023-655-1247
; Sequence 1247, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSER: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1247:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2156 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g292416
; US-09-023-655-1247
Alignment Scores:
Pred. No.: 2,69e-15 Length: 2156
Score: 231.50 Matches: 81

```

Percent Similarity: 44.81%
 Best Local Similarity: 26.30%
 Query Match: 12.37%
 DB: 4
 Gaps: 13

US-09-995-225b-16 (1-353) x US-09-023-655-1247 (1-2156)

QY 22 AlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGlyLeu 41
 Db 153 GCCCTTGGGGCCCACTGCTGCCCCCTGTG--TACTCTGTGATTTGTCATTGGCCGTG 209
 QY 42 ProAlaSerIleLeuThrValIleIleLeuSerGlnLeuValAlaAArgGlnLysSer 61
 Db 210 GTTGGAAACATCTGAGTGTCTGCTGCTTGTGCAATACAAAGAGCTAAATAACATGACC 269
 QY 62 SerTyrSerIleLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPheIleVal 81
 Db 270 AGCATC--TACTCTGAACTGACCTGCCATTTCTGACCTCTTCTGTCAGGCTTCC 326
 QY 82 Phe---ValAspPhe---LeuLeuGluAspPheIleLeuSerGlnMetProGlnVal 99
 Db 327 TTTGATCGACATCAAGTTGAAGATGACTGGGTTTGGTGTGATGCCATG----- 377
 QY 100 ProAspLysIleIleGluValLeuGluPheSerSerIleIleIleThrSerIleTyrIleThr 119
 Db 378 ---TGTAAATCTCTCTCTGCGTTTATTATACAGAGCTGTACAGCGATCTTTTCATC 434
 QY 120 ValProLeuThrIleAspArgTyrIleAlaValCysHis-----ProLeuLysTyr 136
 Db 435 ATCTGCTGACGATTCAGAGTACAGTACCTGGCCATCTGTCACGCCGTGGTTGGCTTCCGGCA 494
 QY 137 HisThrValSerTyrProAlaArgThrArgValIleValSerValTyrIleThrCys 156
 Db 495 CGAGCCGTCACTTTGGTGTATCATCACAGCATCATCATTTGGCCCTGGCATC----- 548
 QY 157 PheLeuThrSerIlePro---TyrTyrTyrTyrProAsnIleTyrThrGluAspTyrIle 175
 Db 549 ---TTGGCTTCAGAACCCAGGCTTATCTTCCAGAACCCAAAGGAAATTCACCTCACAC 605
 QY 176 SerThrSerValHisHis-----ValLeuIleTyrIleHisCysPheThrVal 181
 Db 606 ACCTGACGCTTCACTTCTCTCAGAAAGCTACAGAGTGAAGCTGTTTCAAGCTCTG 665
 QY 182 -----ValLeuIleTyrIleHisCysPheThrVal 191
 Db 666 AAATGAACTCTTGGGCTGTATTCCTTGTGTATCATCTGCTACACA--- 722
 QY 192 TyrLeuValProCysSerIlePhePheIleLeuAsnSerIleIleValTyrIleThr 211
 Db 723 -----GGATTAATTAAGATTCTGTGAAGA 746
 QY 212 ArgLysSerAsnPheArgLeuArgLysTyrSerThrGlyLysThrThrAlaLeuPhe 231
 Db 747 CCAACCAAAATAGAG-----AAGAAATCCAAAGCGTCCGTTGATTTT 788
 QY 232 ThrIleThrSerIlePheAlaThrIleuThrAlaProArgIleIleMetIleLeuTyrHis 251
 Db 789 GTCATCATATCATCTTTTCTCTTGTGACCCCTTCAATTCATCAATTAATTTCT 848
 QY 252 LeuTyrGlyAlaProIle-----GlnAsnArgTyrPheValHisIle 265
 Db 849 GTTTTCAAGACTTCTGTTCACCATGATGTGACAGACAGACATTTG--GACTG 905
 QY 266 MetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPhePheLeuTyr 285
 Db 906 GCTGTGCAAGTACGAGAGTATCCCTAACACGACGCTGTGTCAACCAAGTATCTAC 965
 QY 286 CysPheIleSerLysArgPheArg 293
 Db 966 GCCTTCGTTGGTGAAGGTTCCGG 989

RESULT 11
 US-09-029-027b-1

/ Sequence 1, Application US/09029027B
 / Patent No. 6441133
 / GENERAL INFORMATION:
 / APPLICANT: Walker, Philippe
 / TITLE OF INVENTION: No. 6441133el TRH Receptor
 / FILER REFERENCE: Walker appl.
 / CURRENT APPLICATION NUMBER: US/09/029,027B
 / CURRENT FILING DATE: 1998-01-27
 / NUMBER OF SEQ ID NOS: 5
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 1
 / LENGTH: 1600
 / TYPE: DNA
 / ORGANISM: Rat
 / US-09-029-027b-1
 Alignment Scores:
 Pred. No.: 1,466-14 Length: 1600
 Score: 223.00 Matches: 79
 Percent Similarity: 43.12% Conservative: 62
 Best Local Similarity: 24.16% Mismatches: 97
 Query Match: 11.91% Indels: 90
 DB: 4 Gaps: 13
 US-09-995-225b-16 (1-353) x US-09-029-027b-1 (1-1600)
 QY 31 ValTyrTyrSerLeuLeuLeuCys---LeuGlyLeuProAlaSerIleLeuThrValIle 49
 Db 437 GCTTCTCAGAGCTCTCTGTGTGACACCTGGGCATCGGGCAATGCATGTGATCTGTG 496
 QY 50 IleLeuSerGlnLeuValAlaArgArgGlnLysSerTyrAsn---TyrLeuLeuAla 68
 Db 497 GTGGG-----CTGACCTCAGGATGACATGACACACCACTGCTACTGCTGAGC 550
 QY 69 LeuAlaAlaAlaAspIleLeuValLeuPhePheIleValPheValAspPheLeuGlu 88
 Db 551 CTGGCCCTCGGTGACCTCTCTGCTG----- 577
 QY 89 AspPheIleLeuAsnMetGlnMetProGlnValProAspLysIle----- 103
 Db 578 -----CTGGCTGGGGGTCTGCCAATGTCTGACAGCTTACAGGGGCACTGGATC 628
 QY 104 -----IleGluValLeuGluPheSerIleHisThrSer 115
 Db 629 TATGAGCTGTGCTGCTGCTGGCATCACTTCCAGTACTGGGCAATCAATGTCTCC 668
 QY 116 IleTyrIleThrValProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLys 135
 Db 689 TCCCTTCCATCCCTGAGCTTCACTGTGAGAGGATATATACATTTGCCACCCACTGAGA 748
 QY 136 TyrHisThrValSerTyrProAlaArgThrArgLysValIleValSerValTyrIleThr 155
 Db 749 GCACAGACCGGTGACCTGTGCGCCGGGCCAAACGATCTGTGACAGCATCTGGGGGCTC 808
 QY 156 CysPheLeuThrSerIleProTyrTyrTyr--TyrProAsnIleTyrThrGluAspTyrIle 175
 Db 809 ACGTCCCTTATATGCTCACTGTGCTTCTCTG-----TGATTCGAATGTCCGT 859
 QY 175 eSerThrSerValHisHisValLeuIleTyrIleHis-----CysPheThrValTyr 192
 Db 860 GACAAACAGCGCTTGAATGTGCTTACAA--GTGCCCGGAGACTGTACTGCTGCCATCTGA 918
 QY 192 rLeuValProCysSerIlePhePheIle-----LeuAsnSerIleIleValTyrTyr 209
 Db 919 CTGTGCTGACTTGTGCTGTCTTTTCAATCGAGCCCTTGTGTGACCTGTGCTCTATATG 978
 QY 209 sLeu----- 210
 Db 979 GCTCATCGGAGGATTTTATTTCAGAGCCGTTGTCCAGAAAGCTGTGCAGAGAGAG 1038
 QY 211 -----ArgArgLysSerAsnAsp 216
 Db 1039 GCAGCCCATGGGACAGAGAGCTGCACAGCAAGCACTGCTCAGGGCCAAAGAGC----- 1093

QY 216 eaRgLeuAArgLyTySerThrGlyLysThrAlaIleLeuPheThrIleThrSer11 236
DB 1094 -----TCCAGAAAGCAGGCGCAGAGATGTCGCGCTGTGTGTGCT 1137
QY 236 ePheAlaThrLeuThrAlaPro---ArgIleIleMetIleLeuTyHisLeuTyGlyAl 255
DB 1138 TTTTGCCGCTGTGTGACCCCTTACCGACATGCTGCTCAACTCTTTGTGGCCCA 1197
QY 255 aProIleGlnAsnArgTyrPheuValHisIleMetSerAspIleAlaAsnMetLeuAlaLe 275
DB 1198 GCTTTTCCTGGACCCCTGGCTCTGCTG-----TTCTGCCGCGACCTGTGTCTA 1245
QY 275 uLeuAnThrAlaIleAsnPhePheLeuTyCyPheIleSerLysArgPheArgThrme 295
DB 1246 CACCAACAGCGCTGTCAACCCCTGTCTACACCTGATGTCTCA CAGAAATTCGG----- 1300
QY 295 uAlaAlaAlaThrLeuLys 301
DB 1301 ---GCGGCTTCTCTGAAA 1315
RESULT 12
US-09-919-039-171
Sequence 171, Application US/09919039
Patent No. 6727066
GENERAL INFORMATION:
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
FILE REFERENCE: PA-0035 US
CURRENT APPLICATION NUMBER: US/09/919,039
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 60/222,113
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 401
SOFTWARE: PERL Program
SEQ ID NO 171
LENGTH: 2772
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6727066 331306.1
NAME/KEY: unsure
LOCATION: 2249, 2262
OTHER INFORMATION: a, c, g, or other
US-09-919-039-171
Alignment Scores:
Pred. No.: 3.8e-14 Length: 2772
Score: 222.50 Matches: 81
Percent Similarity: 43.91% Conservative: 56
Best Local Similarity: 25.96% Mismatches: 103
Query Match: 11.89% Indels: 72
DB: 4 Gaps: 13
US-09-995-225b-16 (1-353) x US-09-919-039-171 (1-2772)
QY 22 AlaCyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCyLeuGlyLeu 41
DB 258 GCCTTTGGGGCCCACTGCGCCCTCTCTG---TACTCTTGTGTATTTGCTATTTGGCCCT 314
QY 42 ProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLysSer 61
DB 315 GTTGAACAATCTGTGTGTCTGTCTGTGTCATATACAAAGAGGCTAAATAAATCATGACC 374
QY 62 SerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIleVal 81
DB 375 AGCATC---TACTCTCTGAACCTGTGACCTATTGACCTCTTCTCTGTCAGCTTCCC 431
QY 82 Phe---ValAspPhe---LeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnVal 99
DB 432 TTCTGATGCACTACAAAGTTGAAGATGATGCTGGGTTTGTGTATGCAAG----- 482

QY 100 ProAspLysIleIleGlnValLeuGlnPheSerSerIleHisThrSerIleTrrIleThr 119
DB 483 ---TGTAAATCTCTCTGCTGGTATTATTTATACAGAGGCTTTATACAGAGATCTTTTCATC 539
QY 120 ValProLeuThrIleAspArgTyrIleAlaValCyHis-----ProLeuLysTy 136
DB 540 ATCTGCTACAGATTACAGATGCTGGGCAATGCTCCAGCCGTGTTCGCTGGCGCA 599
QY 137 HisThrValSerTyrProAlaArgThrArgLysValIleValSerValTyrIleThrCy 156
DB 600 CGGACCGTCACTTTTGGTGTCTATCACACAGATCATATTGGGCCCTGGGCATC----- 653
QY 157 PheLeuThrSerIlePro---TyrTyrTrrProAsnIleTrrPheGluLeuTyIle 175
DB 654 ---TTGGCTTCCATGCGAGCTTATCTTTCCAAAGCCAAATGGGAATTCATCTACACAC 710
QY 176 SerThrSerValHis----- 181
DB 711 ACCTGAGCTTCTACTTTCTCTCAGAAACCTACAGAACTGGAAGCTGTTTCAGGCTCTG 770
QY 182 -----ValLeuIleTrrIleHisCyPheThrVal 191
DB 771 AAACGAACTCTTTGGCTGTGATTCCTTTGTTGTGATGATCTATCTGCTACACA--- 827
QY 192 TyrLeuValProCySerIlePhePheIleLeuAsnSerIleValTyrLysLeuArg 211
DB 828 -----GGGATTAATAAGATTCGTAAAG 851
QY 212 ArgLysSerAsnPheArgLeuArgLyTySerThrGlyLysThrAlaIleLeuPhe 231
DB 852 CACCAAAATGAC-----AAGAAATCCAAAGCTCCGTTGATTTT 893
QY 232 ThrIleThrSerIlePheAlaThrLeuThrAlaProArgIleIleMetIleLeuTyHis 251
DB 894 GTCATCATGATCATCTTTTCTCTTTGGACCCCTCACAATTTGATCTATTTCT 953
QY 252 LeuTyGlyAlaProIleGlnAsnArgTrrPheuValHisIleMetSer----- 267
DB 954 GTTTTC-----CAGACTTCTGTTCACCCATGAGTGTAGAGAGACAGACAT 1001
QY 268 -----AspIleAlaAsnMetLeuAlaLeuAnThrAlaIleAsnPh 282
DB 1002 TTGACCTGGCTGTGCAAGTACGAGAGTGTATGCTTACAGACATGCTGTCTAACCC 1061
QY 282 ePheLeuTyCyPheIleSerLysArgPheArg 293
DB 1062 AGTGATCTACGCTTGTGTGTGAGAGTTCCGG 1095
RESULT 13
US-08-676-351-1
Sequence 1, Application US/08676351C
Patent No. 6046026
GENERAL INFORMATION:
APPLICANT: EPFLER, CECIL
APPLICANT: OZEMERGER, BRADLEY
TITLE OF INVENTION: CDNAS ENCODING PROTEINS CLOSELY RELATED
FILE REFERENCE: 0646/1A818US1
CURRENT APPLICATION NUMBER: US/08/676,351C
CURRENT FILING DATE: 1996-09-12
EARLIER APPLICATION NUMBER: PCT/US95/00939
EARLIER FILING DATE: 1995-01-20
EARLIER APPLICATION NUMBER: US 08/185,360
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FaecSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2706
TYPE: DNA
ORGANISM: Rat
US-08-676-351-1

Alignment Scores:

Pred. No.:	6,89e-14	Length:	2706
Score:	220.00	Matches:	75
Percent Similarity:	42.90%	Conservative:	64
Best Local Similarity:	23.15%	Mismatches:	117
Query Match:	11.75%	Indels:	68
DB:	3	Gaps:	13

US-09-995-225B-16 (1-353) x US-08-676-351-1 (1-2706)

```

QY      7 HisleuAlaAlaAsnSerLeuSerTrpSerProGlySerAlaCysGlyLeuGly 26
DB      200 CACCTGCTCTCAATGCTAGTCACAGCCCTTCCTGCCC-----CTTGA 244
QY      27 Phe---ValProValTyrTyrSerLeuLeuCysLeu---GlyLeuProAlaAsn 44
DB      245 CTCAGGTCACCACTGGGGCTTACTTGGCTGTGTCATCGGGGGCTCTCGGGAC 304
QY      45 IleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLeuSerSerTyrAsn 64
DB      305 TGCCCTGTCATGATGTCATC-----CTCAGCACACACCAAGTGAAGACGCTACCAAC 358
QY      65 ---TyrLeuLeuAlaLeuAlaAlaAlaAspIleLeuValLeuPheIleValPhe--- 82
DB      359 ATTACATATTAACTGGACCTGGCTGATACCTGGCTTGTGCTTACACAGCCCTTCAG 418
QY      83 ---ValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValProAsp 101
DB      419 GGCACAGACATCCACTGGCTTGGCTTGGCCATTGGGAATGCACCTC-----TGC 466
QY      102 LysIleIleGlnValLeuGlnPheSerSerIleIleThrSerIleTrpIleThrValPro 121
DB      467 AAGACTGTCTACTGCTACTGACTACTACACATGTTTACGACCTTTACTCGACGCC 526
QY      122 LeuThrIleAspArgTyrIleAlaValCysIleProLeuLysTyrIleThrValSerTyr 141
DB      527 ATGACGCTACAGCGCTATGCGCTATGCGCACTTACCGTCCCTTGATGTTCCGACA 586
QY      142 ProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSerIle 161
DB      587 TCCAGCAAGCCGAGCTGTTAATGCGCATATGGCCCTTGGCTTCAATGGTGGTGT 646
QY      162 ProTyr----- 163
DB      647 CCGTGGCCATCATGGGTTCCAGACAAAGTGAAGATGAAGATGAGCTCGGTGGAG 706
QY      164 -----TyrTrpTrpProAsnIleTrpThrGlnAspTyrIleSerThr 177
DB      707 ATCCCTGCCCCCTCAGGACTATGGGGCCCT----- 736
QY      178 SerValIleIleValLeuIleTrpIleIleCysPheThrValTyrLeuValProCysSer 197
DB      737 ---GTTTGGCCATCTGCACTTCTCTTTCTTC-----ATCATCCCTGTGCTG 784
QY      198 IlePhePheIleLeuAsnSerIleIleValTyrLysLeuArg----- 211
DB      785 ATCATCTCTGCTGTCTACAGCTCATGATGACACACTTGGTGGTCCGCTGCTTCA 844
QY      212 -----ArgLysSerAsnPheArgLeuArgLysTyrSerThrGlyLysThrThrAlaIle 229
DB      845 GGCCTCCCGGAAAGAGCCGAAACCTGGCG-----CGATACATCGACCTG 889
QY      230 LeuPheThrIleThrSerIlePheAlaThrLeuTrpAlaProArgIleIleMetIleLeu 249
DB      890 GTGCTGTAGTGTGGCTGTGTTGTGGCTGTGGCTGTGGAGCTGTGGAGGTTGTGCTG 949
QY      250 TyrHisLeuTyrGlyAlaProIleGlnAsnArgTrpLeuValHisIleMetSerAspIle 269
DB      950 GTTCAAGGACTGGGTGTTCAGCCAGGTAGTGAAGTGAAGTGCATCTCGCGC---TTC 1006
QY      270 AlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPhePheLeuTyrCysPheIleSer 289

```

```

DB      1007 TGCAACGCCCTGCGCTATGTCACAGTGTTCATCCATCTCTATGCTCTGAGAT 1066
QY      290 LysArgPheArg 293
DB      1067 GAGAACTTCAG 1078

```

RESULT 14

US-09-016-434-1446

Sequence 1446, Application US/09016434

Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-young

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

FILING DATE: HEREMITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1446:

SEQUENCE CHARACTERISTICS:

LENGTH: 1317 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: G693907

US-09-016-434-1446

Alignment Scores:

Pred. No.:	2.63e-14	Length:	1317
Score:	219.50	Matches:	72
Percent Similarity:	44.82%	Conservative:	75
Best Local Similarity:	21.95%	Mismatches:	130
Query Match:	11.73%	Indels:	51
DB:	4	Gaps:	11

US-09-995-225B-16 (1-353) x US-09-016-434-1446 (1-1317)

```

QY      19 ProGlySerAlaCysGlyLeuGlyPheValProVal---ValTyrTyrSerLeuLeu 37
DB      165 CCGGGGAGCGCGCGCGCGCGCATGTCGCTATCCAGTCATCTACGCGCTGGTGCG 224
QY      38 CysLeuGlyLeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArg 57
DB      225 CTGTGGGGCTGTGGGAGCAAGCCCTGTCTATCTTCGATCTTCGCG---TACGCCAG 281
QY      58 ArgGlnLysSerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeu 77

```

```

Db      282 ATGAAGACGGCTCAACATCTCTGCTCACTGCGCCCTGACCAAGCTC----- 335
      78 PhePheIleValPheValAspPheLeuGluAspPheIleLeu---AsnMetGlnMet 96
      336 ---TTGATGCTGAGCGCTGCGCTTGAGGCTCTGTCGGCCGCGCTGCGCCACTGGCCCTTC 392
      97 ProGlnValProAspIleIleIleGluValLeuGluPheSerSerIleHisThrSerIle 116
      393 GGCCTCGCTGCTGCGCGGCTGCTCAAGCTGACGGCTCAACAGTTCACACGACGCT 452
      117 TrpIleThrValProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyr 136
      453 TTCTGCTCAACCGCTGCTCAACCGCTGACCGCTGACCGCTGCTGACCGCTGCTGCGCG 512
      137 HisThrValSerTyrProAlaArgThrArgLysValIleValSerValTyrIleThrCys 156
      513 GCGACCTACCGCGCGCGCGCGCGAGCTGACATCACTGGCGCGTGGCTGGCATCC 572
      157 PheLeuThrSerIleProTyr----- 163
      573 CTGTGGTCACTCTCCCATGCGCATCTTGACAGACACGACCGGCTGCGCGCGCGCG 632
      164 -----TyrTrpTrpProAsn---IleTrpThrGluAspTyrIleSerThr 177
      633 GCGGTGGCTGCAACCTGCAAGTGGCCACACCGCGCTGCTGCGCACTTTCGTC----- 686
      178 SerValHisIleValLeuIleTrpIleHisCysPheThrVal---TyrLeuValProCys 196
      687 -----GTCTACACTTTCGCTGCGGCTTCTGCGCGCGCTGCGCGCG 722
      197 SerIlePhePheIleLeuAsnSerIleIleValTyrIleLeuArgArg-----Lys 213
      723 CTGGCCATGCGCTGCTGCTACCTGCTCATCTGCGGAGATGGCGCGCGCGCGCGCG 782
      214 SerAsnPheArgLeuArgLysTyrSerThrGlyLeuThrAlaIleLeuPheThrIle 233
      783 GCTGGCTGGCAGACGGCGCGCGCTGAGAGAAATATCACAGGCTGGCTGAGTGC 842
      234 ThrSerIlePheAlaThrIleuTrpAlaProArgIleIleMetIleLeuTyrHisLeuTyr 253
      843 GTGGTGTCTTGTGCTGCTGCTGAGCTTCTTAGTGCTGACGTGCGAAGCTGCTGCT 902
      254 GlyAlaProIleGlnAsnArgTrpLeuValHisIleMetSerAspIleAlaAsnMetLeu 273
      903 GTGACGAGCGCTTAT-----GCCACCGTCAACACGATGCTCCCTTATCTTC 947
      274 AlaLeuLeuAsnThrAlaIleAsnPhePheLeuTyrCysPheIleSerLysArgPheArg 293
      948 AGCTATGCCAACAGCTGCGCGCAACCTTATCTTATGCTTCTCTCCGACAACTTCCGC 1007
      294 -----ThMetAlaAlaAlaThrLeuLysAlaPhePheLysCys 306
      1008 CGATCTTCCAGCGGGTCTCTGCTGCGCTGCTGCTGCTGGAAGTGTGTAAGTGTCT 1067
      307 GlnLysGlnProValGlnPheTyr 314
      1068 GAGGAGAGAGCGCTGAGTACTACT 1091

```

RESULT 15

```

US-08-149-093A-3
Sequence 3, Application US/08149093A
Patent No. 5658783
GENERAL INFORMATION:
APPLICANT: Bunzow, James R
APPLICANT: Grandy, David K
TITLE OF INVENTION: A No. 5658783el Mammalian Methadone-Specific
TITLE OF INVENTION: Opioid Receptor Gene and Uses
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago

```

```

STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,093A
FILING DATE: 06-NOV-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5658783man, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,311
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..181
FEATURE:
NAME/KEY: CDS
LOCATION: 182..1282
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1283..1452
US-08-149-093A-3
Alignment Scores:
Pred. No.: 3,466-14 Length: 1452
Score: 219.00 Matches: 74
Percent Similarity: 42.90% Conservative: 65
Best Local Similarity: 22.84% Mismatches: 117
Query Match: 11.70% Indels: 68
DB: Gaps: 13
US-09-995-225B-16 (1-353) x US-08-149-093A-3 (1-1452)
      7 HisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGlySerAlaCysGlyLeuGly 26
      275 CACCTGCTCTCCCAAGCTGATGTCACAGCGCTTCTGCGCC-----CTTGA 319
      27 Phe---ValProValValTyrTyrSerLeuLeuCysLeu---GlyLeuProAlaAsn 44
      320 CTCAGGTACCATCTGCGGGGCTCATCTTGGCTGTGTGATGGGGGGCTCTCGGGGAAC 379
      45 IleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLysSerSerTyrAsn 64
      380 TGCCTGTCATGATGTCATC-----CTCAGGACACCCCAAGATGAAAGACACTACCAAC 433
      65 ---TyrLeuLeuAlaLeuAlaAlaAlaPheIleLeuValLeuPhePheIleValPhe--- 82
      434 ATTACATATTTAAATCTGGCTGCTGCTGATACCTGTGCTCTTAACACTGCGCTTCCAG 493
      83 ---ValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValProAsp 101
      494 GCGACAGACATCTACTGCGCTTGGCCATTGGGAAGCACTC-----TGC 541
      102 LysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrValPro 121
      542 AAGACTGTCATGTCATCACTACTACCAACAGTTCACGACCTTACTCTGACCGCC 601
      122 LeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSerTyr 141

```

```
Db      602  ATGAGCGTAGACGGCTATGTGCTATCTGCAACCTTATCCGTGGCCCTTGATGTTGAGACA 661
Qy      142  ProAlaArgThrArgLeuValIleValSerValTyrIleThrCysPheLeuThrSerIle 161
Db      662  TTCACCAAGGCCAGCGTGTATATGTGGCATATGGGCCCTGGCTTCAGTGTTGGTGT 721
Qy      162  ProTyr----- 163
Db      722  CCTGTGGCATCATGGGTTCAGCACAGTGAAGATGAAGATCGAGTGGCTGTGAGAG 781
Qy      164  -----TyrTyrProAsnIleTyrThrGluAspTyrIleSerThr 177
Db      782  ATCCCTGCCCTCAGAGACTATGGGGCCCT----- 811
Qy      178  SerValHisHisValLeuIleTyrIleHisCysPheThrValTyrLeuValProCysSer 197
Db      812  ---GTATTGGCCATCTGCATCTCTCTTTTTCCTTC-----ATCATCCCTGTGCTG 859
Qy      198  IlePhePheIleLeuAsnSerIleIleValTyrIysLeuArg----- 211
Db      860  ATCATCTCTGTCTGTCAAGCCTCATGATTGACGACTTGTGTGTCGCTGCTTCA 919
Qy      212  -----ArgLysSerAsnPheArgLeuArgGlyTyrSerThrGlyTyrThrAlaIle 229
Db      920  GGCTCCCGGAGAGAACCGAACCCTGCG-----CGTATCACTCGACTG 964
Qy      230  LeuPheThrIleThrSerIlePheAlaThrLeuTyrAlaProArgIleIleMetIleLeu 249
Db      965  GTGCTGTGTAGTGTGGCTGTGTGTGTGTGGCTGTGACGCTGTGCAGAGGTGTGTCTG 1024
Qy      250  TyrHisLeuTyrGlyAlaProIleGlnAsnArgTyrPheValHisIleMetSerAspIle 269
Db      1025  GTTCAAGACTGGGTGTTCAGCCAGGTAGTAGAGTGCAGTGCATCTCGGCGC---TTC 1081
Qy      270  AlaAsnMetLeuAlaLeuAsnThrAlaIleAsnPheLeuTyrCysPheIleSer 289
Db      1082  TGCACAGCCCTGGGCTATGTCAACAGTGTGTCAATCCATTCTATGCTTTCCTGAT 1141
Qy      290  LysArgPheArg 293
Db      1142  GAGAACTTCAG 1153
```

Search completed: December 14, 2004, 22:27:16
Job time : 92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2004, 20:35:30 ; Search time 456 Seconds

(without alignments)
4268.142 Million cell updates/sec

Title: US-09-995-225B-16

Perfect score: 1872
Sequence: 1. 1 MHTAHILANSLSKWSPG.....CIKMLVVOYDKXGKPIKVP 353

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.epoc1/US0995225.rnuc.14122004.110734.18062/app.query.fasta_1.519
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blonsum62
-TRANS=humand0.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US0995225.QCGN.1.1.34.qrunac.14122004.110734.18062
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FAPOP=6 -FAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Databases :

Published Applications NA: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1872	100.0	1059	16	US-10-450-590-1	Sequence 1, Appl1
2	1872	100.0	1062	9	US-09-995-225-15	Sequence 15, Appl1
3	1872	100.0	1062	10	US-09-995-225-15	Sequence 15, Appl1
4	1872	100.0	1062	16	US-10-296-2944-3	Sequence 4, Appl1
5	1872	100.0	1062	16	US-10-296-2944-4	Sequence 4, Appl1
6	1872	100.0	1202	14	US-10-094-417-1	Sequence 1, Appl1
7	1872	100.0	2117	17	US-10-779-144-1	Sequence 1, Appl1
8	1868	99.8	1343	10	US-09-813-432-21	Sequence 21, Appl1
9	1868	99.8	1343	15	US-10-174-364-21	Sequence 21, Appl1
10	1868	99.8	1343	16	US-10-246-583-21	Sequence 21, Appl1
11	1868	99.8	1343	17	US-10-689-832-21	Sequence 21, Appl1
12	1857	99.2	1119	10	US-09-813-432-21	Sequence 21, Appl1
13	1857	99.2	1119	10	US-10-174-364-19	Sequence 19, Appl1
14	1857	99.2	1119	16	US-10-246-583-19	Sequence 19, Appl1
15	1857	99.2	1119	17	US-10-689-832-19	Sequence 19, Appl1
16	1857	99.2	1130	16	US-10-333-946-25	Sequence 25, Appl1
17	1857	99.2	2198	15	US-10-314-076-1	Sequence 1, Appl1
18	1844	98.5	1062	14	US-10-012-140-12	Sequence 12, Appl1
19	1844	98.5	1526	14	US-10-012-140-10	Sequence 10, Appl1
20	1841	98.3	1110	15	US-10-314-076-3	Sequence 3, Appl1
21	1841	98.3	2189	14	US-10-219-834-6	Sequence 6, Appl1
22	1756	93.8	1038	14	US-10-094-417-19	Sequence 19, Appl1
23	1747	93.3	1062	16	US-10-450-590-9	Sequence 9, Appl1
24	1643	87.8	957	15	US-10-174-364-84	Sequence 84, Appl1
25	1643	87.8	957	16	US-10-246-583-84	Sequence 84, Appl1
26	1638	87.5	1002	15	US-10-079-384-25	Sequence 25, Appl1
27	1638	87.5	1032	16	US-10-450-590-5	Sequence 5, Appl1
28	1638	87.5	1070	16	US-10-450-590-4	Sequence 4, Appl1
29	1638	87.5	1158	16	US-10-343-650A-21	Sequence 21, Appl1
30	1638	87.5	1826	16	US-10-450-590-6	Sequence 6, Appl1
31	1471	78.6	795	10	US-09-791-932-6	Sequence 6, Appl1
32	1317	70.4	765	15	US-10-328-912-3	Sequence 3, Appl1
33	912	48.7	930	10	US-09-791-932-3	Sequence 3, Appl1
34	742	39.6	1466	16	US-10-328-916-4	Sequence 4, Appl1
35	737	39.4	1125	9	US-09-995-225-5	Sequence 5, Appl1
36	737	39.4	1125	10	US-09-995-225-5	Sequence 5, Appl1
37	737	39.4	1127	14	US-10-094-417-13	Sequence 13, Appl1
38	737	39.4	1330	13	US-10-011-147-1	Sequence 1, Appl1
39	737	39.4	1933	16	US-10-328-916-3	Sequence 3, Appl1
40	737	39.4	1962	16	US-10-333-946-29	Sequence 29, Appl1
41	733	39.2	1125	14	US-10-012-140-15	Sequence 15, Appl1
42	733	39.2	1719	14	US-10-012-140-13	Sequence 13, Appl1
43	668	35.7	34118	15	US-10-017-161-1071	Sequence 1071, Ap
44	668	35.7	34118	15	US-10-293-798-909	Sequence 909, App
45	242	12.9	3364	15	US-10-253-983-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-450-590-1
; Sequence 1, Application US/10450590
; Publication No. US20040076985A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
; FILE REFERENCE: LIO316 Foreign Counties
; CURRENT APPLICATION NUMBER: US/10/450,590
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/254,923
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: US 60/280,110
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/299,474
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1059
; TYPE: DNA

ORGANISM: Homo sapiens
US-10-450-590-1

Alignment Scores:

Pred. No.:	8.7e-178	Length:	1059
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	16	Gaps:	0

US-09-995-225b-16 (1-353) x US-10-450-590-1 (1-1059)

```
QY      1 MetGluH1srThH1sa1aH1s1LeuA1a1aAsnSerSerLeuSerTrpTrpSerProGly 20
DB      1 ATGGAGACACACGACCGCCACCTCGACGCCAACAGCTGGCTGTGGTGTGTCCCGCGC 60
QY      21 SerAlaCySG1yLeuGlyPheValProValValTyTySerLeuLeuLeuCySLeuGly 40
DB      61 TCGGCGCTGGCGCTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY      41 LeuProAlaBn1leuThrVal1le1leuSerG1nLeuVal1a1aArgArgG1nLy 60
DB      121 TTACCCAGAAATATCTTGACAGATGATCTCTCCAGCTGGTGGCAAGAGACAGAAAG 180
QY      61 SerSerTyZAsnTyLeuLeuA1aLeuA1a1aAsp1leuVal1leuPhePhe1le 80
DB      181 TCCCTCCACAAATATCTCTGGACCTCGCTGCGCCGACATCTTGATCTCTTTTCTATA 240
QY      81 ValPheValAsPheLeuLeuG1uAAspPhe1leLeuAsnMetG1nMetProG1nVal1Pro 100
DB      241 GTGTGTGTGGACCTCTGTGTGAAGATTCATCTTGAACATGACATGCTCAGATCCCC 300
QY      101 AspLy1le1leG1uVal1leuG1uPheSerSer1leH1srThSer1leTrp1leThrVal1 120
DB      301 GACAAATCATTAAGTGTGGATTTCTCATTCATCCACACCTCCCATATGATGATCTGTA 360
QY      121 ProLeuThr1leAspArgTy1leA1aVal1CyH1sProlLeuLyTyH1srThVal1Ser 140
DB      361 CCGTTAACCATGTACAGATATATCGCTGTGCGCACCGCTCAAGTACACACGCTCTCA 420
QY      141 TyPProAlAsPheThrArgLyVal1leVal1SerValTy1leThrCySPhelLeuThrSer 160
DB      421 TACCCAGCCCGACCCGGAAGCATGTGAAGTTTACATCATCCGCTCTCTGACAGC 480
QY      161 IleProTyTyTyTrpProAsn1leTrpThG1uAAspTy1leSerThSerVal1H1s 180
DB      481 ATCCCTTATTACGTGTGGCCCAACATCTGACCTGAAGACTACATCAGACCTCTGTGAT 540
QY      181 H1sVal1leu1leTrp1leH1sCySPhenThValTyTyLeuValProCyS1lePhePhe 200
DB      541 CAGCTCTCATCTGATCCACATGCTTCAACCTGCTACCTGCTCCATCTTCTTC 600
QY      201 IleLeuAsnSer1le1leValTyTyTyS1leuA1aG1yS1erAsnPheArgLeuArgLy 220
DB      601 ATCTTGAATCATCATCTGTGTGACAGCTCAGAGAGAAAGCAATTTTGTCTCGTGGC 660
QY      221 TyTySerThGlyLyS1erThThVal1leuPheThr1leThrSer1lePheAlaThrLeu 240
DB      661 TACTCCAGCGGGAAGACACCGCCCATCTTGTTCACCATTAACCTCATCTTGGCCACACTT 720
QY      241 TrpAlaProArg1le1leMet1leLeuTyH1sLeuTyRgLyA1aPro1leG1nAsnArg 260
DB      721 TGGGCCCCCGCATCATCATCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780
QY      261 TrpLeuVal1H1s1leMetSerAsp1leAlaAsnMetLeuAlaLeuLeuAsnThAla1le 280
DB      781 TGGCTGAGACACATCATCTCCAGACATTCGCAACATGCTTACGCCCTTCTTAACACAGCCATTC 840
QY      281 AsnPhePheLeuTyTyCySPhel1leSerLyS1aRgPheArgThrMetVal1a1a1aThrLeu 300
DB      841 AACTTCTTCTCTACTGCTCATCAGCAAGCGGTCCGACACATGGGACCGCCACGCTC 900
```

```
QY      301 TyAlaPhePheLyCySG1nLySG1nProVal1G1nPheTyTrThAsnH1sAsnPheSer 320
DB      901 AAGGCTTTCTTCAAGTGCACAGAGCACTGTACAGTTCTTACACCAATCATTAATCTTTCC 960
QY      321 IleThSerSerProTrp1leSerProAlaAsnSerH1sCyS1leLyS1eLeuValTyR 340
DB      961 ATTAACAAGTGAAGCCCTCGATCTCGCGCGCAACTACACTGCATCAAGATCTGTGTATC 1020
QY      341 G1nTyAspLyAsnG1yLySProl1leLySVal1SerPro 353
DB      1021 CAGTATGACAAAATGGAATAAACCCTATTAAGAATCCCCG 1059
```

RESULT 2

US-09-995-225-15

Sequence 15, Application US/0995225

Publication No. US20020193584A1

GENERAL INFORMATION:

APPLICANT: Chen, Ruoping

APPLICANT: Chu, Zhi Liang

APPLICANT: Dang, Huong T.

APPLICANT: Lowitz, Kevin P.

APPLICANT: Pride, Cameron

TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G

TITLE OF INVENTION: Receptors

FILE REFERENCE: ARBN-0308

CURRENT APPLICATION NUMBER: US/09/995, 225

CURRENT FILING DATE: 2001-11-26

PRIOR APPLICATION NUMBER: 09/170, 496

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: PCT/US99/23938

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: 60/253, 404

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/255, 366

PRIOR FILING DATE: 2000-12-12

PRIOR APPLICATION NUMBER: 60/270, 286

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282, 365

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/270, 266

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282, 032

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282, 358

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282, 356

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/290, 917

PRIOR FILING DATE: 2001-05-14

PRIOR APPLICATION NUMBER: 60/309, 208

PRIOR FILING DATE: 2001-07-31

NUMBER OF SEQ ID NOS: 67

SOFTWARE: PatentIn version 3.1

SEQ ID NO 15

LENGTH: 1062

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: No. US20020193584A1 Sequence

US-09-995-225-15

Alignment Scores:

Pred. No.:	8.74e-178	Length:	1062
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-995-225b-16 (1-353) x US-09-995-225-15 (1-1062)

```
QY      1 MetGluH1srThH1sa1aH1s1LeuA1a1aAsnSerSerLeuSerTrpTrpSerProGly 20
DB      1 ATGGAGACACACGACCGCCACCTCGACGCCAACAGCTGGCTGTGGTGTGTCCCGCGC 60
QY      21 SerAlaCySG1yLeuGlyPheValProValValTyTySerLeuLeuLeuCySLeuGly 40
DB      61 TCGGCGCTGGCGCTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY      41 LeuProAlaBn1leuThrVal1le1leuSerG1nLeuVal1a1aArgArgG1nLy 60
DB      121 TTACCCAGAAATATCTTGACAGATGATCTCTCCAGCTGGTGGCAAGAGACAGAAAG 180
QY      61 SerSerTyZAsnTyLeuLeuA1aLeuA1a1aAsp1leuVal1leuPhePhe1le 80
DB      181 TCCCTCCACAAATATCTCTGGACCTCGCTGCGCCGACATCTTGATCTCTTTTCTATA 240
QY      81 ValPheValAsPheLeuLeuG1uAAspPhe1leLeuAsnMetG1nMetProG1nVal1Pro 100
DB      241 GTGTGTGTGGACCTCTGTGTGAAGATTCATCTTGAACATGACATGCTCAGATCCCC 300
QY      101 AspLy1le1leG1uVal1leuG1uPheSerSer1leH1srThSer1leTrp1leThrVal1 120
DB      301 GACAAATCATTAAGTGTGGATTTCTCATTCATCCACACCTCCCATATGATGATCTGTA 360
QY      121 ProLeuThr1leAspArgTy1leA1aVal1CyH1sProlLeuLyTyH1srThVal1Ser 140
DB      361 CCGTTAACCATGTACAGATATATCGCTGTGCGCACCGCTCAAGTACACACGCTCTCA 420
QY      141 TyPProAlAsPheThrArgLyVal1leVal1SerValTy1leThrCySPhelLeuThrSer 160
DB      421 TACCCAGCCCGACCCGGAAGCATGTGAAGTTTACATCATCCGCTCTCTGACAGC 480
QY      161 IleProTyTyTyTrpProAsn1leTrpThG1uAAspTy1leSerThSerVal1H1s 180
DB      481 ATCCCTTATTACGTGTGGCCCAACATCTGACCTGAAGACTACATCAGACCTCTGTGAT 540
QY      181 H1sVal1leu1leTrp1leH1sCySPhenThValTyTyLeuValProCyS1lePhePhe 200
DB      541 CAGCTCTCATCTGATCCACATGCTTCAACCTGCTACCTGCTCCATCTTCTTC 600
QY      201 IleLeuAsnSer1le1leValTyTyTyS1leuA1aG1yS1erAsnPheArgLeuArgLy 220
DB      601 ATCTTGAATCATCATCTGTGTGACAGCTCAGAGAGAAAGCAATTTTGTCTCGTGGC 660
QY      221 TyTySerThGlyLyS1erThThVal1leuPheThr1leThrSer1lePheAlaThrLeu 240
DB      661 TACTCCAGCGGGAAGACACCGCCCATCTTGTTCACCATTAACCTCATCTTGGCCACACTT 720
QY      241 TrpAlaProArg1le1leMet1leLeuTyH1sLeuTyRgLyA1aPro1leG1nAsnArg 260
DB      721 TGGGCCCCCGCATCATCATCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780
QY      261 TrpLeuVal1H1s1leMetSerAsp1leAlaAsnMetLeuAlaLeuLeuAsnThAla1le 280
DB      781 TGGCTGAGACACATCATCTCCAGACATTCGCAACATGCTTACGCCCTTCTTAACACAGCCATTC 840
QY      281 AsnPhePheLeuTyTyCySPhel1leSerLyS1aRgPheArgThrMetVal1a1a1aThrLeu 300
DB      841 AACTTCTTCTCTACTGCTCATCAGCAAGCGGTCCGACACATGGGACCGCCACGCTC 900
```



```

Db      1 ATGGAGACACGACGCCACCTCGACGACCAAGCTCGCTGTCTTGTTGTCCTCCCGC 60
Qy      21 SerAlaCyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCyLeuGly 40
Db      61 TCGGCTGGGGCTTGGGTTTCGGCCGGCTGCTACTACAGCTCTTGCTGCTCCGCT 120
Qy      41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
Db      121 TTAACGACAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAG 180
Qy      61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
Db      181 TCCTCTTAACAACATATCTTGGCAGCTGGCTGGCCGACATCTTGGCTCTTTTCATA 240
Qy      81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      241 GTGTTTGAGACTTCTGTTGGAAAGATTCACTTGAACATGACAGATGCTTCAGGCTCCC 300
Qy      101 AspValIleIleGluValLeuGluPheSerSerIleHisThrSerIleTyrIleThrVal 120
Db      301 GACAAATCATAGATGCTGGAAATCTCATCTCCATCCACCTCCATATGATTAAGTAT 360
Qy      121 ProLeuThrIleAspArgTyrIleAlaValCyHisProLeuLysTyrHisThrValSer 140
Db      361 CGCTTAACCATTTGACAGTATATCGCTGTGCGCACCCGCTCAAGTACCAACGCTCTCA 420
Qy      141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCyPheLeuThrSer 160
Db      421 TACCCGACCCGCAACCCGAAAGTCATTGTAAGTTTACATCACTCGCTTCGACACAG 480
Qy      161 IleProTyrTyrTrpTrpProAsnIleTyrThrGluAspTyrIleSerThrSerValHis 180
Db      481 ATCCCTTATATCTGGTGGCCCAACATCTGACAGTAAAGTACATACAGACCTCTGGCAT 540
Qy      181 HisValLeuIleTyrIleHisCyPheThrValTyrLeuValProCySerIlePhePhe 200
Db      541 CAGCTCTCATCTGGATCCAGCTGCTTACCGCTTACCTGGTGGCCGCTCATCTTCTTC 600
Qy      201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220
Db      601 ATCTGAACTCATCATTTGTCACAGCTCAGAGAGAGCAATTTTCCTCCCGTGC 660
Qy      221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleHisSerIlePheAlaThrLeu 240
Db      661 TACTCCACGGAGAACCAACCGCATCTGTTACCAATTAACCTCCATCTTGGCACACTT 720
Qy      241 TyrAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValaProIleGlnAsnArg 260
Db      721 TGGGCCCCCGCATCATCATGATTTCTTTCACCTCTATGGGGCGCCCATCCAGAACCCG 780
Qy      261 TyrLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db      781 TGGCTGTGTCACATCATGTCGACATTCGCCAATCATCTGACCTTCTGAAACAGCCATC 840
Qy      281 AsnPhePheLeuTyrCyPhePheIleSerLysArgPheArgThrMetIleAlaAlaThrLeu 300
Db      841 AACTTTCTCTCTACTAGCTTCATCAGCAAGCGGTTCGACCACTGGAGCGGCACAGCTC 900
Qy      301 LysAlaPhePheLysCyGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db      901 AAGGCTTTCTTCAAGGCCAGAACCACTGTAACGTTCTACACCATCATTAATCTTTCC 960
Qy      321 IleThrSerSerProTyrIleSerProAlaAsnSerHisCyAlaLeuMetLeuValTyr 340
Db      961 ATTAACAAGTAGCCCTCGGATCTCGCCGCAAACTCACATGCAAGAAGTGTGGTGTAC 1020
Qy      341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
Db      1021 CAGTATGACAAATAATGAGAAACCTATTAAGTATCCCG 1059

```

RESULT 3
 US-09-995-225-15
 ; Sequence 15, Application US/09995225

```

; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Kuoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pride, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 1998-10-13
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: No. US20030139588A9el Sequence
; US-09-995-225-15

Alignment Scores:
Pred. No.: 8,74e-178 Length: 1062
Score: 1872.00 Matches: 353
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-995-225B-16 (1-353) x US-09-995-225-15 (1-1062)
Qy      1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTyrTrpSerProGly 20
Db      1 ATGGAGACACACGACGCCACCTCGACGACCAAGCTCGCTGTCTTGTTGTCCTCCCGC 60
Qy      21 SerAlaCyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCyLeuGly 40
Db      61 TCGGCTGGGGCTTGGGTTTCGGCCGGCTGCTACTACAGCTCTTGCTGCTCCGCT 120
Qy      41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
Db      121 TTAACGACAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAG 180
Qy      61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
Db      181 TCCTCTTAACAACATATCTTGGCAGCTGGCTGGCCGACATCTTGGCTCTTTTCATA 240
Qy      81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100

```

```

Db      241 GGTGTTGGAGCTTCTGTGAGAGATTCTTGAACATGCAAGATGCTTACGGTCCC 300
QY      101 AspylslellegluValleuGlupheserlellesthrserleletrpIlethrVal 120
Db      301 GACAAGATCATGAAGTGTGGAAATTCATCATCACTCCATATGATTAAGTCTTA 360
QY      121 ProleuthrIleAspArgTyrIleAlaValCysHisProleuValGlyHisThrValSer 140
Db      361 CCGTTAACCATGACAGGTAATTCGCTGTGCAACCCGCTCAAGTACACACAGGCTCA 420
QY      141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db      421 TACCCAGCCCGCACCCGGAAGTCAATTGAAAGTTTACATCACTGCTTCTTACACAGC 480
QY      161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db      481 ATCCCTTTACTGGTGGCCCAACATCTGACCTGAGACTGAAAGTACATACACACCTGTGCAT 540
QY      181 HisValIleuIletrpIleHisCysPheThrValTyrIleuValProCysSerIlePhePhe 200
Db      541 CACGCTCATCTGGATCCATCGCTTACCGCTTACCTGGTGGCTTCTCATCTTCTTC 600
QY      201 IleuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db      601 ACTTGAACCTAACTATTGTTGTACAGCTCAGAGAGAGCAATTTCGTCCTCCGAGC 660
QY      221 TyserThrGlyLysThrThrAlaIleLeuPheThrIlePheSerIlePheAlaThrIleu 240
Db      661 TACTCCAGGGGAGAGACACCGCATTTGTTCACATACCTTCATCTTGGCACACTT 720
QY      241 ThrAlaProArgIleIleMetIleLeuTyrHisLeuTyrValIleProIleGlnAsnArg 260
Db      721 TGGGCCCCCGCAATCATATGATTTCTTTACACCTGATGGGGCGCCATCCAGAACCG 780
QY      261 TrpleuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db      781 TGGCTGGTGCATCATATGTCGACATGTCACACATGTCAGCCCTTGAACACAGCCATC 840
QY      281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrIleu 300
Db      841 AACTCTTCTCTACTGCTTCATCAGCAGCGGTTCGACACATGACACCCGACAGCTTC 900
QY      301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db      901 AAGGCTTTCTTCAAGTGCCAGAGCAACCTGTACAGTTCTACACCATATCACTTTTCC 960
QY      321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleuMetLeuValTyr 340
Db      961 ATAAACAAGTAGCCCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGTGATC 1020
QY      341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
Db      1021 CAGTATGACAAAATGGAATAACCTATAAAGTATCCCCG 1059

```

```

RESULT 4
US-10-296-294A-3
; Sequence 3, Application US/10296294A
; Publication No. US20040029224A1
; GENERAL INFORMATION:

```

```

; APPLICANT: TERAO, Yasuko
; APPLICANT: MATSUI, Hideki
; APPLICANT: SHINTANI, Yasushi
; TITLE OF INVENTION: No. US20040029224A1el G Protein-Coupled Receptor and its DNA
; FILE REFERENCE: 2734 USOP
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/10/296,294A
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/JP01/04643
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: JP 2000-170446
; PRIOR FILING DATE: 2000-06-23

```

```

; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 3
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Human
US-10-296-294A-3

Alignment Scores:
Pred. No.:      8 74e-178      Length:      1062
Score:          1872.00      Matches:      353
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels: 0
DB:               16          Gaps: 0

US-09-995-225B-16 (1-353) x US-10-296-294A-3 (1-1062)

QY      1 MetGluHisThrHisAlaHisLeuAlaIleAsnSerSerLeuSerTrpTrpSerProGly 20
Db      1 ATGGAGCACACGACAGCCCACTCGCAGACCAACAGCTCGTGTGTTGGTGGTCCCCGGC 60
QY      21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db      61 TGGGCTGGGCTTGGGTTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
QY      41 LeuProAlaAsnIleLeuThrValIleIleuSerGlnLeuValAlaArgArgGlyLys 60
Db      121 TTACACGACAAATATCTTACAGTATCATCTCTCCAGCTGGTGGGAGAGACAGAG 180
QY      61 SerSerTyrAsnTyrIleuValAlaAlaAlaAspIleLeuValIleuPhePheIle 80
Db      181 TCTCTTCAACATATCTTGTGGACTCGCTGTCGACATCTTGGTCTCTTTTCATA 240
QY      81 ValPheValAspPheLeuLeuGluAspPheIleuAsnMetGlnMetProGlnValPro 100
Db      241 GGTGTTGGAGCTTCTGTGAGAGATTTCATCTTGAACATGACAGATGCTCAGTCCC 300
QY      101 AspylslellegluValleuGlupheserlellesthrserleletrpIlethrVal 120
Db      301 GACAAGATCATGAAGTGTGGAAATTCATCATCACTCCATATGATTAAGTCTTA 360
QY      121 ProleuthrIleAspArgTyrIleAlaValCysHisProleuValGlyHisThrValSer 140
Db      361 CCGTTAACCATGACAGGTAATTCGCTGTGCAACCCGCTCAAGTACACACAGGCTCA 420
QY      141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db      421 TACCCAGCCCGCACCCGGAAGTCAATTGAAAGTTTACATCACTGCTTCTTACACAGC 480
QY      161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db      481 ATCCCTTTACTGGTGGCCCAACATCTGACCTGAGACTGAAAGTACATACACACCTGTGCAT 540
QY      181 HisValIleuIletrpIleHisCysPheThrValTyrIleuValProCysSerIlePhePhe 200
Db      541 CACGCTCATCTGGATCCATCGCTTACCGCTTACCTGGTGGCTTCTCATCTTCTTC 600
QY      201 IleuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db      601 ACTTGAACCTAACTATTGTTGTACAGCTCAGAGAGCAAAATTTCGTCCTCCGAGC 660
QY      221 TyserThrGlyLysThrThrAlaIleLeuPheThrIlePheSerIlePheAlaThrIleu 240
Db      661 TACTCCAGGGGAGAGACACCGCATTTGTTCACATTACTTCATCTTGGCACACTT 720
QY      241 ThrAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db      721 TGGGCCCCCGCATCATATGATTTCTTACACCTGTAAAGGGGCGCATCCAGAACCGC 780
QY      261 TrpleuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db      781 TGGCTGGTGCATCATATGTCGACATGTCACACATGTCAGCCCTTCTGAAACACAGCCATC 840

```

QY	281	IsnHsPheLeuNYrCySpheNlleSerValArgPhaArgHrMetLalaIaIaThrLeu	300
Db	841	AACTCTTCCTCTCACTGCTTCATCAGCAAGCGGATTCGCGACATGGACGCGCAAGCTTC	900
QY	301	LysAlaPhePheLysCySGlnLysGlnProValGlnPheYrThrAsnHisAsnPheSer	320
Db	901	AAGCTCTTCTTCAAGTCCAGAAAGCAACCTGTACAGTTCACACCAATCATTAATTTTCC	960
QY	321	IleHrSerSerProTrpIleSerProAlaAsnSerHisCysAlleLysMetLeuValYr	340
Db	961	ATAACAAGTAGCCCTCGATCTCGCCGGCAAACTCACACGTGCATCAAGATGCTGGGTAC	1020
QY	341	GlnYrAspLysAsnGlyLysProIleLysValSerPro	353
Db	1021	CAGTATGACAAAATGGAAAACCTATTAAAGTATCCCCG	1059

RESULT 5
 US-10-296-294A-4
 ; Sequence 4, Application US/10296294A
 ; Publication No. US20040029224A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TERAO, Yaeko
 ; APPLICANT: MATSUI, Hiideki
 ; APPLICANT: SHINTANI, Yasuichi
 ; TITLE OF INVENTION: No. US20040029224A1e1 G Protein-Coupled Receptor and its DNA

```

: CURRENT APPLICATION NUMBER: US/10/296,294A
: CURRENT FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: PCT/JP01/04643
: PRIOR FILING DATE: 2001-06-01
: PRIOR APPLICATION NUMBER: JP 2000-170446
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: JP 2000-194926
: PRIOR FILING DATE: 2000-06-23
: NUMBER OF SEQ ID NOS: 8
: SEQ ID NO 4
: LENGTH: 1062
: TYPE: DNA
: ORGANISM: Human
: US-10-296-294A-4

```

Alignment Scores:	
Pred. No.:	8,74e-178
Score:	1872.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	16
US-09-995-225B-16 (1-353) x US-10-296-294A-4 (1-1062)	
	Length: 1062
	Matches: 353
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

[illegible]

Db	301	GACAAAGATCARGAAGAGTGGATTTCTCATCCATCCACACCTCCATATGAGTTAACTGTA	360
Oy	121	ProLeuthrIleapAgtYrIleAlaValCyshiSproLeuYsYrHiSerValSer	140
Db	361	CCGTTAACCATTTGACAGGTATATATGCTGTCTGCACCCGCTCAAGTACCAACAGGCTCTCA	420
Oy	141	TyrTroAlaAaGthraGlyValIleValSerValYrIleThrCySphLeuThrSer	160
Db	421	TACCAAGCCCGACCCGGAAAGTAAAGTGGTTTAAACATCACACTGCTCTTCTACACAGC	480
Oy	161	IleProYrYrYrTTPProAsnIleThrThrGluAspYrIleSerThrSerValHis	180
Db	481	ATCCCTTATCTGTGTGTCGCCCAACATCTGGACTAAAGCTACATCAGACCTCTGTGCAT	540
Oy	181	HisValLeuIleTrrIleHisCySphThrValYrLeuValProCySerIlePhePhe	200
Db	541	CAGCTCCTCATCTGGATCCACTGTTCACCGTTCACCTGGAGCCCTGCTCATCTTCTTC	600
Oy	201	IleLeuAsnSerIleIleValYrYrIleLeuArgArgYsSerAsnPheArgLeuArgIly	220
Db	601	ATCTTGAACCTCAATCATTTGGTATACACTCAGGAGGAAGCAATTTCTGCTCCGTGGC	660
Oy	221	TyrSerThrGlyYsThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu	240
Db	661	TACTCCAGGGGAAAGACACCGGCATCTTGTTCACATTACTCATCTTTGGCCACACTT	720
Oy	241	TrrAlaProArgIleIleMetIleLeuYrHisLeuYrGlyAlaProIleGlnAsnArg	260
Db	721	TGGGCCCCCGGCATCATCATGATTTCTTTACACTCTATGGGGGCCCATCCAGAACGGC	780
Oy	261	TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle	280
Db	781	TGGCTGGGACATCATCATGTCCGACATTTGCCAACATGTAGCCCTTCGAAACACGCCATTC	840
Oy	281	AsnPhePheLeuYrCySphIleSerYrAspPheArgThrMetAlaAlaIleThrLeu	300
Db	841	AACCTCTTCTCTACTGCTCATATGAGAAAGGGTTCGACACATGAGGACCGCCACGCTC	900
Oy	301	LysAlaPhePheYsCySgInlySgInProValGlnPheYrThrAsnHisAsnPheSer	320
Db	901	AAGGCTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTACACCAATCAATACTTTTCC	960
Oy	321	IleThrSerSerProTrrIleSerProAlaAsnSerHisCySgIleYsMetLeuValYr	340
Db	961	ATAACAGATGACCCCTGATCTCCCGGCAAACTCACATGCAATCAAGATGCTGTGTAC	1020
Oy	341	GlnTrrAspYsAsnGlyYsProIleIleValSerPro	353
Db	1021	CAGTATGACAAATAATGAAAAACCTATTAAGATATCCCCG	1059

```

RESULT 6
US-10-094-417-1
; Sequence 1, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jisang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030045685A1e1 Receptor
; FILE REFERENCE: 018781-008110US
; CURRENT APPLICATION NUMBER: US/10/094,417
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/802,803
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/276,649
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1202
; TYPE: DNA

```



```

Db      580 GACAGATCATAGAGTGTGAAATTCATCATCCACACCTCATATGATTAAGTGA 639
Qy      121 Proleuthrillepharygryrillealavalcyahisproleuylsyrhistrvalser 140
Db      640 CCGTAAACCATTAACAGATATACGCTGTGACCCCGTCAAGTACCAACAGGCTCA 699
Qy      141 TTPROALARGTHIRGLYSVALILEVALSERVALTYRILETHRCYPHELEUTHSER 160
Db      700 TACCCAGCCCGACCCCGAAGTCAATTTGTAAGTGTTCATCACTGCTCTCCGACACG 759
Qy      161 ILEPROTYRTYRTPTRPROASNIETPTTHGLUAPRYRIIESERTHSERVALHIS 180
Db      760 ATCCCTATTAATCGTGGGCCCAACATCTGAGTGAAGACTACATCAGCACTCTGTGCAT 819
Qy      181 HISEVALLEUIETPTTHIRGLYSYSPHEETHVALTYRILEUVALPROCYSSERTILEPHE 200
Db      820 CAGGCTCATCTGATTCACCTGCTTCAACGCTACCTGTGCTCCCTCATCTCTTC 879
Qy      201 ILEUASNSERTILEVALTYRILYSLEUARGLYSERASNPHEARGLEUARGLY 220
Db      880 ATCTTAACCTCATCTGATTCATGATTCAGTCAAGTCAAGAGAGCAATTTTCGTCCTG 939
Qy      221 TYSERTHRGLYSRTHRALALEUPTHERTHRIETHSERILEPHEALATHRLAU 240
Db      940 TACTCCACGGGGAAGACACCGCCATCTGTTCACCATTACTCCATCTTGCCACACTT 999
Qy      241 TTPALAPROARGILELLEWECTILEUITYRHLSEUITYRGLVALAPROLLEGINANAG 260
Db      1000 TGGGCCCGCCGATCATCATGATTCATTACACCTGATGGGCGCCATCCAGAAACCG 1059
Qy      261 TTPLEUVALHIELEWECTSERAPILALEASNMELUVALLEUANTHRLAILE 280
Db      1060 TGGCTGTGATCATCTGATTCATGTCGACATTCGCCACATGTCACCTTGTGAACAGACCAT 1119
Qy      281 ASNPHEPHELEUITYRCSYPHEILESERLYSARGPHEARGTHRMEALALATHRLAU 300
Db      1120 AACTTCTTCTCTACAGCTTCAACAGACGCTTCCGACCATGGAGCGCCACAGCTC 1179
Qy      301 LYSALAPHEPHELYSCYSGINLYSGINPROVALGLPHEITYRTHRANHISANPHESE 320
Db      1180 AAGGCTTTCTTCAAGTGCAGAGCAACCTGACAGTTTCAACCATCACTTAATCTTTC 1239
Qy      321 ILETHSERSEPRORTPILESERPROALASNSERTHISCYRIELYSMELEUVALTYR 340
Db      1240 ATACAGAGTACCCCTGGAGTCTGCGCGCAACTCAACATCAAGATGTGTGTAC 1299
Qy      341 GINTYASPLYASNGLYLSPROIILEYVALSERPRO 353
Db      1300 CAGTATGACAAATAATGAAAACTATTAAGTATCCCGC 1338

RESULT 8
US-09-813-432-21
; Sequence 21, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
; APPLICANT: Taupler Jr., Raymond J
; APPLICANT: Majumder, Kamud
; APPLICANT: Spaderma, Steven K
; APPLICANT: Smithson, Glenda
; APPLICANT: Mezes, Peter S
; APPLICANT: Vermet, Corine A. M.
; TITLE OR INVENTION: NO. US20030148485A1el polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729
; CURRENT APPLICATION NUMBER: US/09/813,432
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22

```

```

; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/193,843
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-813-432-21

Alignment Scores:
Pred. No.: 3,14e-177 Length: 1343
Score: 1668.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Beech Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
Gaps: 0

US-09-995-225b-16 (1-353) x US-09-813-432-21 (1-1343)

Qy      1 MetGLHITHTHIAHISLEUALAIAANSERLEUSERTPTTPSERPROGLY 20
Db      2 ATGGAGCACACGACCCCACTCGCAGCAACAGCTCCGTCCTTGCTGCTCCCGGC 61
Qy      21 SerAlaCyGlyLLeuGlyPheValProValValTYRTHYRSELEUENUCYSLANGLY 40
Db      62 TCGGCTGGGCTGGGTTCGTCGCGGTGCTACTACAGCTCTTGCTGCTGCTCGGT 121
Qy      41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGln 60
Db      122 TTAACGACAAATATCTTGACAGTATCATCTCCAGCTGGTGCAAGAGACGAAG 181
Qy      61 SerSerTYRAsnTYRLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
Db      182 TCCTCCTAACATATCTCTTGACACTCGCTGCGACATCATCTGCTCTTTTCATA 241
Qy      81 ValPheValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      242 GTGTTGTGACTCTCTGTGGAAGATTTGATCTTGAAATGACAGATGCTCAGGTCGCC 301
Qy      101 AspLYIleIleGluValLeuGlnPheSerSerIleIleTHSERIleTPHleTHVal 120
Db      302 GACAGATCATAGAGTGTGAAATTCATTCATCAATCCACACTCCATATGATTAAGTGA 361
Qy      121 Proleuthrillepharygryrillealavalcyahisproleuylsyrhistrvalser 140
Db      362 CCGTAAACCATTAACAGATATACGCTGTGACCCCGTCAAGTACCAACAGGCTCA 421
Qy      141 TTPROALARGTHIRGLYSVALILEVALSERVALTYRILETHRCYPHELEUTHSER 160
Db      422 TACCCAGCCCGACCCCGAAGTCAATTTGTAAGTGTTCATCACTGCTCTCCGACACG 481
Qy      161 ILEPROTYRTYRTPTRPROASNIETPTTHGLUAPRYRIIESERTHSERVALHIS 180
Db      482 ATCCCTATTAATCGTGGGCCCAACATCTGAGTGAAGACTACATCAGCACTCTGTGCAT 541
Qy      181 HISEVALLEUIETPTTHIRGLYSYSPHEETHVALTYRILEUVALPROCYSSERTILEPHE 200
Db      542 CAGGCTCATCTGATTCACCTGCTTCAACGCTACCTGTGCTCCCTCATCTCTTC 601
Qy      201 ILEUASNSERTILEVALTYRILYSLEUARGLYSERASNPHEARGLEUARGLY 220

```

```

Db      602 ATCTGAACCTCATCTATGCTGTCAGAGGAGAGAGCAATTTTCGTCTCCGTGGC 661
Qy      221 TyrSerThGlyLysThrThraAlaIleuPheThrIleThSerIlePheAlaThrIleu 240
Db      662 TACTCCACGGGAGAGCCACCGCATCTGTTCCACATTACTCCATCTTGGCCACACTT 721
Qy      241 TrpAlaProArgIleIleMetIleLeuTyrlsleuTyrglyAlaProIleGlnAsnArg 260
Db      722 TGGGCCCCCGCATCATCTGATTTCTTTACCACTCTAAGGGGCCCATCCAGAACCGC 781
Qy      261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db      782 TGGCTGTGACACATCATGCTCCAGCATGCGCAACATGACCTTCTTAACACAGCCATC 841
Qy      281 AsnPhePheLeuTyrlsCysPheIleSerIleAspArgPheArgPheMetAlaAlaIleu 300
Db      842 AACTCTTCTCTTACTGCTTCACTGACGAGCGGTTCCGACCATGCGCAGCCGCGCTC 901
Qy      301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrlsThrAsnHisAsnPheSer 320
Db      902 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTAAGTTCTACACCAATCAATAACTTTTCC 961
Qy      321 IleThSerSerProThrIleSerProAlaAsnSerHisCysIleIleuMetLeuValTyr 340
Db      962 ATACAGATAGCCCTGGATCTCGCCGCAAACTCACACATGCACTGCAAGATGCTGTGTAC 1021
Qy      341 GlnTyAspLysAsnGlyLysProIleLysValSerPro 353
Db      1022 CAGTATGACAAATAATGMAAACCTATAAAGTATCCCCG 1060

```

RESULT 9

```

US-10-174-364-21
; Sequence 21. Application US/10174364
; Publication No. US20030216308A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2
; CURRENT APPLICATION NUMBER: US/10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2) .. (1060)
US-10-174-364-21

```

Alignment Scores:

```

Pred. No.: 3,14e-177 Length: 1343
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 15 Gaps: 0

```

US-09-995-225b-16 (1-353) x US-10-174-364-21 (1-1343)

```

Qy      1 MetGlnHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpTrpSerProGly 20
Db      2 ATGAGACACACGACCGCCACTCGACGCAACACTGCTGTCTGGTGTGTCCTCCCGC 61
Qy      21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCysLeuGly 40
Db      62 TGGCTGTGAGCTTGGGTTTCGTGCCCGGTCTACTACAGCTTCTTGTGCTGTGCTGGT 121
Qy      41 LeuProAlaAsnIleLeuThrValIleIleuSerGlnLeuValAlaArgGlnLys 60
Db      122 TTACGAGCAATATCTTGACAGTATCATCTCTCCAGCTGTGTGGCAAGAACAGAG 181
Qy      61 SerSerTyrrAsnTyrlsLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db      182 TCCTCTCAACTATCTTGTGCACTGCTGCTGCCACATCTTGCTCTTTTCATA 241
Qy      81 ValPheValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      242 GTGTTTGTGACTTCCTGTTGGAAGATTTTATCTTGAACAGAGATGCTCAGGTCCC 301
Qy      101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db      302 GACAAAGATCATGAGAGGCTGGATTTCTCATCTCATCCACACTCCCATATGATTACTGTA 361
Qy      121 ProLeuThrIleAspArgTyrlsAlaValCysHisProLeuLysTyrlsHisThrValSer 140
Db      362 CCGTTAACATTAAGACATATATCATCTGTGACCCGCTCAAGTACCAACAGGCTCTCA 421
Qy      141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db      422 TACCCAGCCCGCACCCCGAAGATCATGTGAGGTTTATCATCACTGCTCTCCAGCAC 481
Qy      161 IleProTyrlsTrpTrpProAsnIleTyrlsGlnAspTyrlsSerThrSerValHis 180
Db      482 ATCCCATTAATACGTGTGGCCCAACATCTGACATGAGATCATCACTGCTGTGAT 541
Qy      181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db      542 CAGTCTCATCTGATCCACATGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
Qy      201 IleLeuAsnSerIleIleValTyrlsLeuAlaGlyLysSerAsnPheArgLeuArgIly 220
Db      602 ATCTTGAACCTCATATCTGTGACAGCTCAGAGAGAGCAATTTTCGTCTCGTGGC 661
Qy      221 TyrSerThGlyLysThrThraAlaIleuPheThrIleThSerIlePheAlaThrIleu 240
Db      662 TACTCCACGGGAGAGCCACCGCATCTGTTCCACATTACTCCATCTTGGCCACACTT 721
Qy      241 TrpAlaProArgIleIleMetIleLeuTyrlsleuTyrglyAlaProIleGlnAsnArg 260
Db      722 TGGGCCCCCGCATCATCTGATTTCTTTACCACTCTAAGGGGCCCATCCAGAACCGC 781
Qy      261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db      782 TGGCTGTGACACATCATGCTCCAGCATGCGCAACATGACCTTCTTAACACAGCCATC 841
Qy      281 AsnPhePheLeuTyrlsCysPheIleSerIleAspArgPheArgPheMetAlaAlaIleu 300
Db      842 AACTCTTCTCTTACTGCTTCACTGACGAGCGGTTCCGACCATGCGCAGCCGCGCTC 901
Qy      301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrlsThrAsnHisAsnPheSer 320
Db      902 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTAAGTTCTACACCAATCAATAACTTTTCC 961

```

QY 321 ILeThrSerSerProTPrIleSerProAlaAaSerHisCyseTlleLeuValTyr 340
Db 962 ATACAGAGTACCCCTGGATCTCGCCGCAACTCAGACATGCAAGATGGGTGTAC 1021

QY 341 GlnTYrAspLysAsnGlyLysProIleLeuValSerPro 353
Db 1022 CAGTATGACAAAATGGAAGAACTATATAAGTATCCCCG 1060

RESULT 10
US-10-246-583-21
; Sequence 21, Application US/10246583
; Publication No. US20040058862A1
; GENERAL INFORMATION:
; APPLICANT: Majumder
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1060)
US-10-246-583-21

Alignment Scores:
Pred. No.: 3,14e-177 Length: 1343
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 16 Gaps: 0

US-09-995-225b-16 (1-353) x US-10-246-583-21 (1-1343)

QY 1 MetGluHisThrIleAlaHisLeuAlaAlaAsnSerSerLeuSerTyrTyrSerProGly 20
Db 2 ATGGAACACACGACCGCCACCTCGACGACCAAGCTCGTGTCTGGTGTCTCCCGGC 61

QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuValGly 40
Db 62 TCGCGCTGGCGCTTGGGTTTCGCGCCGCTGCTTACTACAGCCTTTCGCTGCTCGGT 121

QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
Db 122 TTACCGAGCAAAATATCTTGAACAGTATCATCTCTCCACGCTGGCGAAGACAGAG 181

QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db 182 TCCTCTTACAACTATATCTTGGCACTCGCTGCGCCGACATCTTGGTCTCTTTTTCATA 241

QY 81 ValPheValAspPheLeuLeuGlyLysAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 242 GTGTTTGTGGACTTCTCGTTGGAAATTTTCATCTTGAACATGCAAGATGCTCGATCCC 301

QY 101 AspLysIleIleGlnValLeuGlyPheSerSerIleHisThrSerIleTyrIleThrVal 120
Db 302 GACCAAGATCATAGAAAGTGTGCAATTCATCATCATCACAACCTCCATATGAGATTACTGTA 361

QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 362 CCGTTAACATGATGACAGGTATACCTGTCTGCAACCGCTCAAGTACACACAGGTTCTCA 421

QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 422 TACCCAGCCCGACCCGGAAGTATGTAAGTTTATCATCATCCTGCTTCGACAGC 481

QY 161 IleProTyrTyrTyrTyrProAsnIleTyrThrGlnAspTyrIleSerThrSerValHis 180
Db 482 ATCCCTTATTAAGTGTGGCCCAACATCTGAGCTGAAGACTACATCAGCATCTGTGTCAT 541

QY 181 HisValLeuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 542 CACGTCTCATCTGGAATCCATGCTTACCGCTTACCTGCTGCTTCATCTTCTTC 601

QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLysArgLys 220
Db 602 ATCTTAACATCATCATTTGTGTACAGCTCAGAGGAAGACAAATTTTGTCTCCGTGG 661

QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 662 TACTCCACGCGGAAGACCAACCGCATCTGTTCACCATTAACCTCCATCTTGTGCACACTT 721

QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 722 TGGGCCCCCGCATCATCATGATCTTTCACACCTTATGAGGCGGCCATTCAGAAACGGC 781

QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 782 TGGCTGTACATCATCATGTCGACATGTCACATGTCATGCTTCGAAACAGCAGCATC 841

QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 842 AACTCTCTCTCATCTGCTTCATCAGCAAGCGGTTCCGCAACATGCGACGCGCACGCTC 901

QY 301 LysAlaPhePheLysGlyGlnArgGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 902 AAGGCTTTCTTCAAGTCCAGGAAGCAACCTGTACAGTTCTACACCAATCATTAATCTTTC 961

QY 321 ILeThrSerSerProTPrIleSerProAlaAaSerHisCyseTlleLeuValTyr 340
Db 962 ATACAGAGTACCCCTGGATCTCGCCGCAACTCAGACATGCAAGATGGGTGTAC 1021

QY 341 GlnTYrAspLysAsnGlyLysProIleLeuValSerPro 353
Db 1022 CAGTATGACAAAATGGAAGAACTATATAAGTATCCCCG 1060

RESULT 11
US-10-689-832-21
; Sequence 21, Application US/10689832
; Publication No. US20040121360A1
; GENERAL INFORMATION:
; APPLICANT: Majumder, Kamud
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND AMINO ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729D1V1
; CURRENT APPLICATION NUMBER: US/10/689,832
; CURRENT FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835

Score: No.:	3,146,77	Length:	1343
Percent Similarity:	1868.00	Matches:	352
Best Local Similarity:	99.72%	Conservative:	1
Query Match:	99.79%	Mismatches:	0
DB:	17	Indels:	0
		Gaps:	0

Db	542	CACGCTCTCATGTGGATCCACTGCTTACCGGTCACCTGGGCGCTGCTCCATCTTCTTC	601
QY	201	lLeuAaenSerileleValTyrLysAuaArgArgLysSerAaPheArgLeuArgLg	220
Db	602	ATCTGAACATCAATCATGTGTACAAAGCTCAGAGAGAAAGCAATTTTCGTCTCCGTGGC	661
QY	221	TyrSerThrGlyLysThrThraAlaIleuAenPheThrIleThrSerIlePheAlaThrLeu	240
Db	662	TACTCCAGGGAGAACCAACCGCCCATCTGTTCACCATTAATCTTCATCTTGGCCACACTT	721
QY	241	TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAaArg	260
Db	722	TGGGCCCCCGCCATCATCATGATTCCTTTACCACTCTATGGGGGGCCCATCCAGAACCGC	781
QY	261	TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAaenThrAlaIle	280
Db	782	TGGTGGTACATCATCACTGTCGACATTCGCAACATGCTAGGCCCTTTCGAAACACAGCCATC	841
QY	281	AsnPhePheLeuTyrCysPheIleSerLysAaArgPheArgThrMetAlaAlaAlaThrLeu	300
Db	842	AACCTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGACACCAATGCGACGCCCACTC	901
QY	301	LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAaPheSer	320
Db	902	AAAGCTTTCTTCAGATCCAGAACCAACCTGTACAGTTTACCACTCAATCACTTTTCC	961
QY	321	IleThrSerSerProTyrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr	340
Db	962	ATTAACAAGTACGCCCTCGAGTCTCGCCGCGCAAACTCACACATGATCAAGATGGCTGTAC	1021
QY	341	GlnTyrAspLysAsnGlyLysProIleLysValSerPro	353
Db	1022	CAGTATGACAAATAATGAAAACCTATTAAGTATCCCGC	1060


```

Db      61 TGGCGCTGGCGCTGGCTTCTGCTCCCGTGGTCTACAGCTCTTCTGCTGGCTCGGT 120
QY      41 LeuProAlaAsnIleLeuThrValIleIleuSerGlnLeuValAlaArgGlnIlys 60
Db      121 TTACGAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAG 180
QY      61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db      181 TCTCTCTACACTATCTTCTGGCACTCGCTGCGCAGCTCTGGTCTCTTTTTCATA 240
QY      81 ValPheValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      241 GTGTTGTGACTCTCTGTGGAAGATTTCATTCTTGAACAGCAAGTCCCTCAGAGTCCC 300
QY      101 AspValIleIleGlnValLeuGlnPheSerSerIleIleThrSerIleTyrIleThrVal 120
Db      301 GACAGATCATAGAAAGTGTGAATTCATCATCATCCACCTCCATATGATTACTGTA 360
QY      121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db      361 CCGTTAACATTGACAGGTATATGCTGTCTGCCACCCGCTCAAGTACCAACGGTCTCA 420
QY      141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db      421 TACCAAGCCGCGACCCGGAAGATCATGTGAAGTTTATCATCACTGCTCTCGACCAAGC 480
QY      161 IleProTyrTyrTyrTyrProAsnIleTyrThrGlnAspTyrIleSerThrSerValHis 180
Db      481 ATCCCTATTAAGTGTGGCCCAACATCTGAGTGAAGACATCAATCAAGCACTCTGTGCAT 540
QY      181 HisValLeuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db      541 CAGCTCTCATCTGATCCACTGCTTCAACCGTCTACCTGGTGGCCGCTCCATCTTCTTC 600
QY      201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgLys 220
Db      601 ATCTTGACTCATCATTTGTGTACAAAGCTCAGAGAGAGAGCAATTTTGTCTCCGTGCGC 660
QY      221 TyrSerThrGlySerThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db      661 TACTCCACGGGGAGAGACACCGCCATCTGTTTACCATTAACCTCCATCTTGGCACACT 720
QY      241 TrpAlaProArgGlyIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db      721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGGGGCCGCCATCCAGAACCGC 780
QY      261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db      781 TGGCTGTACATCATATGTCGACATGSCCAATGCTAGCCCTCTGAAACACAGCCATC 840
QY      281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaIleThrLeu 300
Db      841 AACTCTTCTCTCTCATCTGCTTCAATCAGCAAGCGGTCCGACCATGGACGCGCCACGCTC 900
QY      301 LysAlaPhePheLysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db      901 AAGGCTTTTCAATGTCAGAGACCACTGTACAGTCTTCAACCAATCATTAATCTTTTCC 960
QY      321 IleThrSerSerProTyrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db      961 ATTAACAGTAGCCCTGGATCTCGCCGCAAACTCAGACTGCATCAAGATCTGCTGTATC 1020
QY      341 GlnTyrAspLysAsnGlyLysProIleLys 350
Db      1021 CAGTATGCAAAATATGAAAACCTATATAAA 1050

```

RESULT 14
 US-10-246-583-19
 ; Sequence 19, Application US/10246583
 ; Publication No. US20040058862A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Majumder

```

; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1116)
; US-10-246-583-19

Alignment Scores:
Pred. No.: 3,03e-176 Length: 1119
Score: 1857.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.20% Indels: 0
DB: 16 Gaps: 0

US-09-995-225b-16 (1-353) x US-10-246-583-19 (1-1119)
QY      1 MetGlnHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTyrTrpSerProGly 20
Db      1 ATGGAGCAGACAGCAGCGCCACCTCGCAGCCAAAGCTCGCTGTGCTGGTCCCGGC 60
QY      21 SerAlaGlyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
Db      61 TCGGCGTGGCGCTGGGTGCTTTCGCGCCGCTGCTTACTACAGCTCTTGTGCTGGCTCGGT 120
QY      41 LeuProAlaAsnIleLeuThrValIleIleuSerGlnLeuValAlaArgGlnIlys 60
Db      121 TTACGAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAG 180
QY      61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db      181 TCTCTCTACACTATCTTCTGGCACTCGCTGCGCAGCTCTTGTGCTCTTTTTCATA 240
QY      81 ValPheValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      241 GTGTTGTGACTCTCTGTGGAAGATTTCATTCTTGAACAGCAAGTCCCTCAGAGTCCC 300
QY      101 AspValIleIleGlnValLeuGlnPheSerSerIleIleThrSerIleTyrIleThrVal 120
Db      301 GACAGATCATAGAAAGTGTGAATTCATCATCATCCACCTCCATATGATTACTGTA 360
QY      121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db      361 CCGTTAACATTGACAGGTATATGCTGTCTGCCACCCGCTCAAGTACCAACGGTCTCA 420

```

QY 141 TyrProAlaArgThrArglyValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGCAGCCGGAAAGCATGTGTAAGTTTACATCACCTGCTTCGAGACGAC 480
 QY 161 IleProTyrTyrTrpTrpProAenIleTrpThrGluAapTyrIleSerThrSerValHis 180
 DB 481 ATCCCTATTACTGAGTGGGCCCAACATCTGAGATGAGACTACATACAGACCTCTGTCAT 540
 QY 181 HisValIleuIleTrpIleHisCysPheThrValTyrIleuValProCysSerIlePhePhe 200
 DB 541 CAGGCTCTCATCTGATCCAGCTGCTTCACCGCTACCTGAGCCCTGCTCATCTTCTTC 600
 QY 201 IleLeuAenSerIleIleValTyrIlyleuArgArglySerAenPheArgIleuArgly 220
 DB 601 ATCTTGAACTCAATCATTTGTGTCAAGCTCAGAGAGAAAGCAATTTTCTCTCCGTCG 660
 QY 221 TyrSerThrGlyIlyerThrAlaIleLeuPheThrIleThrSerIlePheAlaThrIleu 240
 DB 661 TACTCCAGGGGAGAGACCCGCACTCTTTCACCACTTACCTCCATCTTTCGACACTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisIleuTyrGlyAlaProIleGlnAenArg 260
 DB 721 TGGGCCCCCGCATCATCATGATTTCTTTACACCTGATGGGGGCCCATCCAGAACCGC 780
 QY 261 TrpLeuValHisIleMetSerAapIleAlaAenMetLeuAlaLeuAenThrAlaIle 280
 DB 781 TGGCTGTACATCATATGTCCGACATTCGCAACATGTCTACCCCTTCGAAACAGCCATC 840
 QY 281 AenPhePheLeuTyrCysPheIleSerIleArgPheArgThrMetAlaAlaThrIleu 300
 DB 841 AACTTTCTCTCTACAGCTTCATCAGACAGCGGTTCGCAACATGGACGCGCCAGCTC 900
 QY 301 LysAlaPhePheLeuCysGlnLysGlnProValGlnPheTyrThrAenHisAenPheSer 320
 DB 901 AAGGCTTTCTTCAAGGCCAGAAAGCAACCTGTACAGTTCTACACCATCATTTCTTTC 960
 QY 321 IleThrSerSerTrpTrpIleSerProAlaAenSerHisCysIleLysMetLeuValTyr 340
 DB 961 ATTAACAAGTAGCCCTGGATCTGCGCGCAACTCACATGATCAAGATGCTGTGTATC 1020
 QY 341 GlnTyrAspLysAenGlyLysProIleLys 350
 DB 1021 CAGTATGACAAATAATGAAAACTTATMAA 1050
 RESULT 15
 US-10-689-832-19
 ; Sequence 19, Application US/10689832
 ; Publication No. US20040121380A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Majmuder, Kamud
 ; TITLE OF INVENTION: Novel Polypeptides and Amino Acid Encoding Same
 ; FILE REFERENCE: 15966-729DIv1
 ; CURRENT APPLICATION NUMBER: US/10/689,832
 ; CURRENT FILING DATE: 2003-10-20
 ; PRIOR APPLICATION NUMBER: 09/813,432
 ; PRIOR FILING DATE: 2001-03-20
 ; PRIOR APPLICATION NUMBER: 60/190,835
 ; PRIOR FILING DATE: 2000-03-20
 ; PRIOR APPLICATION NUMBER: 60/190,768
 ; PRIOR FILING DATE: 2000-03-20
 ; PRIOR APPLICATION NUMBER: 60/190,972
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 60/191,199
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 60/191,947
 ; PRIOR FILING DATE: 2000-03-24
 ; PRIOR APPLICATION NUMBER: 60/192,665
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192,657
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192,984
 ; PRIOR FILING DATE: 2000-03-28

; PRIOR APPLICATION NUMBER: 60/192,664
 ; PRIOR FILING DATE: 2000-03-28
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 19
 ; LENGTH: 1119
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-689-832-19
 Alignment Scores:
 Pred. No.: 3 03e-176 Length: 1119
 Score: 1657.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.20% Indels: 0
 DB: 17 Gaps: 0
 US-09-995-225B-16 (1-353) x US-10-689-832-19 (1-1119)
 QY 1 MetGluHisThrHisAlaHisIleuAlaAenSerSerLeuSerTrpTrpSerProGly 20
 DB 1 ATGAGACACACGCGACGCCACCTCGCAGCCACACAGCTCCGTGTGTCGTCGCCGCGC 60
 QY 21 SerAlaCysGlyIleuGlyPheValProValValTyrTyrSerLeuLeuGlyLeuGly 40
 DB 61 TCGGCTGGGGCTGGGTTCGTCGCGCGTGTCTTACACAGCTCTTGTGTCGTCGTCG 120
 QY 41 LeuProAlaAenIleuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
 DB 121 TTACGACAAATATCTTGAACATGATCATCTCTCCAGCTGTGGCAAGAACAGAGAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAapIleuValIleuPhePheIle 80
 DB 181 TCCTCTACAAATATCTTGGCATCTGGCATCGCTGCGGAGCATCTGTGCTCTTTTCARA 240
 QY 81 ValPheValAapPheLeuLeuGlyAapPheIleLeuAenMetGlnMetProGlnValPro 100
 DB 241 GTGTGTGTGACTCTCTGTGTGAAGATTCATCTTGAACATGAGATGCTCAGTCCCGC 300
 QY 101 AapLysIleIleGluValIleuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 301 GACAAATCATAGAAATGCTCGAATTTCTATCATCATCACCTCCATATGATTAATCTGA 360
 QY 121 ProLeuThrIleAapArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
 DB 361 CGCTTAACCATTGAACAGTATATCGCTGTCTGCAACCGCTCAAGTACACAGAGCTCA 420
 QY 141 TyrProAlaArgThrArglyValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGCAGCCGGAAAGCATGTGTAAGTTTACATCACCTGCTTCGAGACG 480
 QY 161 IleProTyrTyrTrpTrpProAenIleTrpThrGluAapTyrIleSerThrSerValHis 180
 DB 481 ATCCCTATTACTGAGTGGGCCCAACATCTGAGATGAGACTACATACAGACCTCTGTCAT 540
 QY 181 HisValIleuIleTrpIleHisCysPheThrValTyrIleuValProCysSerIlePhePhe 200
 DB 541 CAGGCTCTCATCTGATCCAGCTGCTTCACCGCTACCTGAGCCCTGCTCATCTTCTTC 600
 QY 201 IleLeuAenSerIleIleValTyrIlyleuArgArglySerAenPheArgIleuArgly 220
 DB 601 ATCTTGAACTCAATCATTTGTGTCAAGCTCAGAGAGAAAGCAATTTTCTCTCCGTCG 660
 QY 221 TyrSerThrGlyIlyerThrAlaIleLeuPheThrIleThrSerIlePheAlaThrIleu 240
 DB 661 TACTCCAGGGGAGAGACCCGCACTCTTTCACCACTTACCTCCATCTTTCGACACTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisIleuTyrGlyAlaProIleGlnAenArg 260
 DB 721 TGGGCCCCCGCATCATCATGATTTCTTTACACCTGATGGGGGCCCATCCAGAACCGC 780

```
Oy 261 TTPLeuValHleIleMerSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
    |||||
Db 781 TGGCTGGTACACATCATGTCGACATGCGCAACATGCTAGGCCCTTGTGACACAGCCATC 840
Oy 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
    |||||
Db 841 AACCTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGACCATGCGAGCCGCAAGCTC 900
Oy 301 LysAlaPhePheLeuCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
    |||||
Db 901 AAGGCTTCTTCTCAAGTGCAGAGACACCTGTACAGTCTACACCAATCATTAACCTTTCC 960
Oy 321 IleThrSerSerProTyrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
    |||||
Db 961 ATTAACAAGTAGCCCTGGAATCTCGCCGCGCAAACTCACTGCATCAAGATGCTGTGTAC 1020
Oy 341 GlnTyrAspLysaengLysProIleLys 350
    |||||
Db 1021 CAGTATGACAAAAATGAAAAAAGCTATATAAA 1050
```

Search completed: December 14, 2004, 22:25:44
Job time : 465 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:04:39 ; Search time 15.4373 Seconds

(Without alignments)
2200.157 Million cell updates/sec

Title: US-09-995-225B-16

Perfect score: 1872

Sequence: 1 MHTAHLAANSLSMWSPG.....CIKMLVYQYDKXGKPIKVP 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

283416

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	242	12.9	411	2	I56444
2	242	12.9	412	2	S23436
3	241	12.9	393	2	A39251
4	238.5	12.7	398	2	JN0708
5	231.5	12.4	355	2	A45177
6	226	12.1	352	2	JE0296
7	225	12.0	357	2	UC7319
8	219.5	11.7	388	2	JN0605
9	219	11.7	384	2	A47249
10	218	11.6	423	2	UC7677
11	217.5	11.6	340	2	TI8704
12	217.5	11.6	394	2	UC7209
13	217.5	11.6	658	2	UC8011
14	217	11.6	367	2	JC2421
15	217	11.6	367	2	JC2421
16	217	11.6	367	2	JC2421
17	214.5	11.5	595	2	JC8012
18	213.5	11.4	355	2	I49339
19	213	11.4	352	2	A45747
20	213	11.4	370	2	S43087
21	213	11.4	384	2	JC4629
22	212	11.3	360	2	T23063
23	211.5	11.3	376	2	T24368
24	211.5	11.3	380	2	JC2338
25	211.5	11.3	504	2	T29338
26	210.5	11.2	380	2	S36143
27	209	11.2	424	2	JH0164
28	208.5	11.1	363	2	I57940
29	207.5	11.1	399	2	A46632

30	207.5	11.1	428	2	JN0692	cholecystokinin ty
31	206.5	11.0	380	2	A48227	kappa opioid recep
32	206.5	11.0	380	2	UC2434	kappa opioid recep
33	206.5	11.0	391	2	A41795	somatostatin recep
34	206.5	11.0	391	2	C41795	somatostatin recep
35	205.5	11.0	353	2	UC2492	G protein-coupled
36	205.5	11.0	391	2	A39297	somatostatin recep
37	205	11.0	355	2	G02436	chemokine (C-C) re
38	205	11.0	355	2	UC5067	G protein-coupled
39	205	11.0	418	2	S59601	vasopressin recep
40	204.5	10.9	352	2	S60024	bradykinin Bi rece
41	204.5	10.9	380	2	A55259	kappa opioid recep
42	204	10.9	353	2	S28787	neuropeptide y/pep
43	204	10.9	372	2	I36532	delta opioid recep
44	203.5	10.9	314	2	TI5831	hypothetical prote
45	203.5	10.9	477	2	UC7913	capsa receptor (CGI

ALIGNMENTS

RESULT 1

I56444

Chytrotophin-releasing hormone receptor - mouse

C:Species: Mus sp. (mouse)

C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 11-Jan-2000

C/Accession: I56444

R/Seller, R.E.; Taylor, P.L.; Lamb, R.F.; Zabavnik, J.; Anderson, L.; Eldre, K.A.

J. Mol. Endocrinol. 10, 199-206, 1993

A>Title: Functional expression and molecular characterization of the chytrotophin-relea

A/Reference number: I56444; PMID:93249585; PMID:8387312

A/Accession: I56444

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-411 <RES>

C/Cross-references: GB:S60053; NID:G300151; PIDN:AA26491.1; PID:G300152

C/Superfamily: adenosine receptor A1

Query Match 12.9%; Score 242; DB 2; Length 411;

Best Local Similarity 22.6%; Pred. No. 1.1e-12;

Matches 83; Conservative 71; Mismatches 125; Indels 88; Gaps 14;

QY	19	PGSACGLGF-VPVVYVSYLLC-IGLPANITVYIISQVARRQKSSYN-VYLALAAADLT	75
DB	16	POVAVALLEYOVVITLLVAVICGIVGNTIMVAVVNR-TKHMRTATNCYLSAVADLM	73
QY	76	VLFFIVFDLLEDFTLNMQMPQVDPKI-----TEVLEFSSIHSTIWITVPL	122
DB	74	VL-----VAAGLPNITDSIGSWYGVGCLCTIYLOYLGNASSCITAF	119
QY	123	TIDRYAVCHPLKHTVSYSPARTKVIYVSVITCFITSIPIYVW-----PNITW-EDYIS	176
DB	120	TIERYAIGHPIKAGFLCTFSPRAKIIIFVMA-----FTSIYCMIMFLDLNISTYDAIV	176
QY	177	TSVHVLWTHCTVYLVPCSIFFILNISTYVVL-----	210
DB	177	ISCGYKISRNYSPILMDPGVFPVPMPLATVLVGFARILFLNIPDDPKENSTTKN	236
QY	211	---RRKSNFRLG-----YSTGKTALIFTTISIPATLMAP-RITMILVHLGAPI	257
DB	237	DSHQKNNKNNLWTRNCFSTVSSRKQVTKMLAVVILFALLMMPRLTVVNSFLSSP	296
QY	258	QNRWLVIHNSDIANMLALINTAINFLLYCFISKRFPTMAATLKAEFKOKQVOPQYTNH	317
DB	297	QENWFLL-----FCRCIYVNSAINPVIYVIMSQKFR-----AAFRKLCNCKQKPTERANV	348
QY	318	NFSITSS 324	
DB	349	SVALNTS 355	

RESULT 2

S23436

thyroliberin receptor - rat
 N/Alternate names: thyrotropin-releasing hormone receptor
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C/Accession: S23436; I53279; A49168; PQ0326
 R/de la Pena, P.; Delgado, L.M.; del Camino, D.; Barros, F.
 Biochem. J. 284, 891-899, 1992
 A/Title: Cloning and expression of the thyrotropin-releasing hormone receptor from GH(3)
 A/Reference number: 823436; MUID:92322017; PMID:11377915
 A/Accession: S23436
 A/Molecule type: mRNA
 A/Residues: 1-412 <PEN>
 A/Cross-references: UNIPROT:Q01717; EMBL:X64630; NID:957394; PIDN:CAA5913.1; PID:957395
 R/Kimura, N.; Arai, K.; Sahara, Y.; Suzuki, H.; Kimura, N.
 Endocrinology 134, 432-440, 1994
 A/Title: Estradiol transcriptionally and posttranscriptionally up-regulates thyrotropin-
 A/Reference number: I53279; MUID:94102223; PMID:8215956
 A/Accession: I53279
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-412 <RES>
 A/Cross-references: GB:D17469; NID:g464199; PIDN:BA04289.1; PID:g464200
 R/Zhao, D.; Yang, J.; Jones, K.E.; Gerald, C.; Suzuki, Y.; Hogan, P.G.; Chin, W.W.; Tash
 Endocrinology 130, 3529-3536, 1992
 A/Title: Molecular cloning of a complementary deoxyribonucleic acid encoding the thyroci
 A/Reference number: A49168; MUID:92283212; PMID:1317787
 A/Accession: A49168
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 112,'D',14-290,'K',292-412 <ZHA>
 A/Experimental source: GH cells
 A/Note: sequence extracted from NCBI backbone (NCBIN:104788, NCBI:P104795)
 R/Yamada, M.; Mouden, T.; Satoh, T.; Iizuka, M.; Murakami, M.; Iriuchijima, T.; Mori, M.
 Biochem. Biophys. Res. Commun. 184, 367-372, 1992
 A/Title: Differential regulation of thyrotropin-releasing hormone receptor mRNA levels b
 A/Reference number: PQ0326; MUID:92231953; PMID:1373613
 A/Accession: PQ0326
 A/Molecule type: mRNA
 A/Residues: 30-58,'P',60-222,'T',224-261 <YAM>
 A/Experimental source: strain Wister
 A/Note: the authors translated the codon ACA for residue 88 as Ala
 C/Superfamily: adenosine receptor A1
 C/Keywords: G protein-coupled receptor; transmembrane protein
 F1-22/Domain: transmembrane #status predicted <TM1>
 F1-24/Domain: transmembrane #status predicted <TM2>
 F1-92/Domain: transmembrane #status predicted <TM3>
 F116-140/Domain: transmembrane #status predicted <TM4>
 F165-186/Domain: transmembrane #status predicted <TM5>

Query Match 12.9%; Score 242; DB 2; Length 412;
 Best Local Similarity 22.6%; Pred. No. 1.1e-12;
 Matches 83; Conservative 71; Mismatches 125; Indels 88; Gaps 14;

QY 19 PSACGLGF-VPVVVYSLLC-IGLPANILTYIISQVAROKSSYN-YLALAAADIL 75
 DB 16 POVAVALERQVVTILLVITIGLGVGNIMVAVLVMR--TKMRPTNICYLVSIAVADIL 73
 QY 76 VLFFIVFDLLEDFILNMQMPQVDPDKI-----IEVLEFSSIHSTIWTIVPL 122
 DB 74 VL-----VAAGLPNITDSIGSWVYGVCGLCTITLYQYIGINASSCSITAF 119
 QY 123 TIDRIYAVCHPLKHTVSPARTKRVYSVYTCTFLTSIPYYW-----PNIMT-EDYIS 176
 DB 120 TIERYIALCHPIKAQPLCTFSRAKKIILFWMA---FTSIYCMLEFLDLNISTYKDAIV 176
 QY 177 TSVHVLVIMHCFVYLVPCSIFFILNSIIVKL----- 210
 DB 177 ISCGKISRNYSPYILMDPGVFVEMILATVLYGFLARILFLNPISDPKSKMKN 236
 QY 211 ---RRKSNFRLG-----YSTGKTTALLFTTISIFATLMAP-RIMILYHLYGAPI 257
 DB 237 DSTHONKMINLATNTRCFNSTVSSRKQVTKLAAVVVILFALLMMPYRILVAVNSLSSPF 296

QY 258 QNRMLVHMSDIAMMLALNTALINFFLYCFEISKEPRTMAATLKAFFKQKOPVOFTYTH 317
 DB 297 QENWFLL-----PCRICIYINSAINPVIYIMLSQKFR-----AAFRKLCNCKQKPTERKANY 348
 QY 318 NFSITSS 324
 DB 349 SVALNYS 355

RESULT 3
 A39251
 thyrotropin-releasing hormone receptor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004
 C/Accession: A39251
 R/Straub, R.E.; Frech, G.C.; Joho, R.H.; Gershengorn, M.C.
 Proc. Natl. Acad. Sci. U.S.A. 87, 9514-9518, 1990
 A/Title: Expression cloning of a cDNA encoding the mouse pituitary thyrotropin-releasin
 A/Reference number: A39251; MUID:91088548; PMID:2175902
 A/Accession: A39251
 A/Molecule type: mRNA
 A/Residues: 1-393 <STR>
 A/Cross-references: UNIPROT:P21761; GB:M59811; GB:M37490; NID:g202153; PIDN:AAA40480.1;
 C/Superfamily: adenosine receptor A1
 C/Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 12.9%; Score 241; DB 2; Length 393;
 Best Local Similarity 22.1%; Pred. No. 1.3e-12;
 Matches 81; Conservative 75; Mismatches 123; Indels 88; Gaps 15;

QY 19 PSACGLGF-VPVVVYSLLC-IGLPANILTYIISQVAROKSSYN-YLALAAADIL 75
 DB 16 POVAVALERQVVTILLVITIGLGVGNIMVAVLVMR--TKMRPTNICYLVSIAVADIL 73
 QY 76 VLFFIVFDLLEDFILNMQMPQVDPDKI-----IEVLEFSSIHSTIWTIVPL 122
 DB 74 VL-----VAAGLPNITDSIGSWVYGVCGLCTITLYQYIGINASSCSITAF 119
 QY 123 TIDRIYAVCHPLKHTVSPARTKRVYSVYTCTFLTSIPYYW-----PNIMT-EDYIS 176
 DB 120 TIERYIALCHPIKAQPLCTFSRAKKIILFWMA---FTSIYCMLEFLDLNISTYKDAIV 176
 QY 177 TSVHVLVIMHCFVYLVPCSIFFILNSI---IYV----- 208
 DB 177 VSCGKISRNYSPYILMDPGVFVEMILATVLYGFLARILFLNPISDPKSKMKN 236
 QY 209 -KLRRKSNFRLG-----YSTGKTTALLFTTISIFATLMAP-RIMILYHLYGAPI 257
 DB 237 DSIHONKMINLATNTRCFNSTVSSRKQVTKLAAVVVILFALLMMPYRILVAVNSFLSSPF 296
 QY 258 QNRMLVHMSDIAMMLALNTALINFFLYCFEISKEPRTMAATLKAFFKQKOPVOFTYTH 317
 DB 297 QENWFLL-----PCRICIYINSAINPVIYIMLSQKFR-----AAFRKLCNCKQKPTERKANY 348

QY 318 NFSITSS 324
 DB 349 SVALNYS 355

RESULT 4
 JN0708
 thyrotropin-releasing hormone receptor - human
 C/Species: Homo sapiens (man)
 C/Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
 C/Accession: S40682; JN0759; S50151; S50152; I38356; JN0708
 R/Matre, V.; Karlsson, H.B.; Wright, M.S.; Lundell, I.; Fjeldheim, A.K.; Gabrielsen, O.S.
 Biochem. Biophys. Res. Commun. 195, 179-185, 1993
 A/Title: Molecular cloning of a functional human thyrotropin-releasing hormone receptor.
 A/Reference number: S40682; MUID:93371401; PMID:835824
 A/Accession: S40682
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-398 <MAT>

A:Cross-references: UNIPROT:P34981; EMBL:X75071; NID:G404157; PIDN:CAA52965.1; PID:G4041
R.Yamada, M.; Monden, T.; Satch, T.; Satch, N.; Murakami, M.; Iituchijima, T.; Kakegawa,
Biochem. Biophys. Res. Commun. 199, 737-745, 1993
A:Title: Pituitary adenomas of patients with acromegaly express thyrotropin-releasing ho
rmon gene.
A:Reference number: JN0759; MUID:93384596; PMID:8396925
A:Accession: JN0759
A:Molecule type: mRNA
A:Residues: 1-398 <YAM>
A:Cross-references: GB:D16845; NID:G577631; PIDN:BA04120.1; PID:G577632
R.Hinuma, S.; Hasegawa, M.; Ogi, K.; Tanaka, H.; Nagai, Y.; Onda, H.
Biochim. Biophys. Acta 1219, 251-259, 1994
A:Title: Molecular cloning and functional expression of a human thyrotropin-releasing ho
rmon gene.
A:Reference number: S50151; MUID:95002135; PMID:7918619
A:Accession: S50151
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-263 <HIN>
A:Accession: S50152
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 267-398 <H12>
R.Duthie, S.M.; Taylor, P.L.; Anderson, L.; Cook, J.; Eldne, K.A.
Mol. Cell. Endocrinol. 95, R11-R15, 1993
A:Title: Cloning and functional characterization of the human TRH receptor.
A:Reference number: I38356; MUID:94063224; PMID:8243797
A:Accession: I38356
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-398 <RES>
A:Cross-references: EMBL:X72089; NID:G440155; PIDN:CAA50979.1; PID:G440156
C:Genetics:
A:Gene: GDB:TRHR
A:Cross-references: GDB:228955; OMIM:188545
A:Map position: 8q23-8q23
C:Superfamily: adenocine receptor A1
C:Keyword: G protein-coupled receptor; receptor; transmembrane protein
F:29-51/Domain: transmembrane #status predicted <TM1>
F:62-83/Domain: transmembrane #status predicted <TM2>
F:101-121/Domain: transmembrane #status predicted <TM3>
F:146-168/Domain: transmembrane #status predicted <TM4>
F:194-215/Domain: transmembrane #status predicted <TM5>
F:267-288/Domain: transmembrane #status predicted <TM6>
F:297-319/Domain: transmembrane #status predicted <TM7>
Query Match 12.7%; Score 238.5; DB 2; Length 398;
Best Local Similarity 22.4%; Pred. No. 2.1e-12;
Matches 80; Conservative 70; Mismatches 120; Indels 87; Gaps 13;
QY 28 VPPVYVYSLLC-IGLPANILTVIILSQLVARROKSSYN-YLLAALAAADIVLFYFVDP 85
DB 26 VVILVILVLIIGIGVIGNIVLVVNR--TKHRTPTNCLVSLAVALDMLV----- 75
QY 86 LLEDPLINQMOPVPPKI-----IEVLEPSSIHSTIWTPLTIDRYAVCH 132
DB 76 -----VAAGLPITISIGSMWYGVGCLCTIYLOVGINASSCSITATIRYAIICH 129
QY 133 PLKHYTVSPARTRKIVSVYITCPLTSPYVW-----PNWT-EDYISTSVHVLWI 186
DB 130 PIKAGLCTFSRAKKIIFVMA---PTSLYCMMLPFLDLNISTYDAIVISGKYSRN 186
QY 187 HCFYVYLVCSIFPLINSIIVYCL-----RRKSN-- 215
DB 187 YSPYIYLVDFGVYVPMILATVLPRIARILFLNPIPSDPKNSKTKWDSHTQNTN 246
QY 216 -----FLRGSSTGKTAILEFTISIFATLWAP-RIMILVHLGAPIONLWINS 267
DB 247 VNTSRNCFSTYSRRQVTKMLAVIYILFALLMPRTLVVNVSPSSPQEMWFL--- 303
QY 268 DIANMLALNTAINFLYLCISRRFRMAATLKAPKCKOPVQFYTNHFSITSS 324
DB 304 -FCRICIYINSAINPIVYIMLSOKFR---AARFKLCNCKOKPTEKPNAYSVALNYS 355

RESULT 5
A45177
Chomokine (C-C) receptor 1 - human
N/Alternate names: C-C CKR-1; macrophage inflammatory protein-1-alpha receptor
C/Species: Homo sapiens (man)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #ext_change 09-Jul-2004
C/Accession: A45177; I5671
R/Moore, K.; Dighegorio, D.; Mak, J.Y.; Horuk, R.; Schall, T.J.
Cell 72, 415-425, 1993
A:Title: Molecular cloning, functional expression, and signaling characteristics of a C
C chemokine receptor.
A:Reference number: A45177; MUID:93161416; PMID:7679328
A:Accession: A45177
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-355 <NEO>
A:Cross-references: UNIPROT:P32246; GB:L10918; NID:G292416; PIDN:AAA36543.1; PID:G29241
A:Experimental source: HL60 cells
A>Note: sequence extracted from NCBI backbone (NCBI:P.124876)
R/Gao, J.
J. Exp. Med. 177, 1421-1427, 1993
A:Title: Structure and functional expression of the human macrophage inflammatory 1 alpi
A:Reference number: I55671; MUID:93240122; PMID:7683036
A:Accession: I55671
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: GB:L10918; NID:G292416; PIDN:AAA36543.1; PID:G292417
C:Genetics:
A:Gene: GDB:CMKRL1; CMKR-1
A:Cross-references: GDB:138446; OMIM:601159
A:Map position: 3p21-3p21
C:Superfamily: vertebrate rhodopsin
C:Keyword: disulfide bond; G protein-coupled receptor; glycoprotein; phosphoprotein; t
F:36-60/Domain: transmembrane #status predicted <TM1>
F:71-91/Domain: transmembrane #status predicted <TM2>
F:108-129/Domain: transmembrane #status predicted <TM3>
F:147-171/Domain: transmembrane #status predicted <TM4>
F:205-223/Domain: transmembrane #status predicted <TM5>
F:240-264/Domain: transmembrane #status predicted <TM6>
P:288-305/Domain: transmembrane #status predicted <TM7>
P:5/Binding site: carbohydrate (Asn) (covalent) #status predicted
P:24-273, 106-183/Disulfide bonds: #status predicted
P:345/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted
Query Match 12.4%; Score 231.5; DB 2; Length 355;
Best Local Similarity 26.3%; Pred. No. 7.3e-12;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;
QY 22 AGCGRPVYVYSLLCGLPANIITVILSQLVARROKSSYNYLLAALAAADIVLPFY 81
DB 31 AFGAQLPLPL-YSLVFVIGLVGNILVVLVYVYKRLKMTSL-YLLMLAISDLPLFTLP 88
QY 82 F-VDF-LLEDPLINQMOPVPPKIIEVLEPSSIHSTIWTPLTIDRYAVCH---PLKY 136
DB 89 FVIDYKLDWDVFGDAM---CKILSGFYTGYSIFPIILTIDRYALVHAVALNA 144
QY 137 HTVSPARTRKIVSVYITCPLTSP-YYWPNWITEDYISTSVH----- 181
DB 145 RVTYFVYISIIIMALAI---LASMPGLVFSKTYQEFTHTCSLHPHPSLRWKLFGAL 201
QY 182 -----VLWVHCFVYLVPGSIFPLINSIIVYKLRKSNRPLRGYSTGKTAIIF 231
DB 202 KNLPLGLVPLPLVMIICYT-----GIKILLRPNE-----KSKAVRLIF 242
QY 232 TTTSIPATLWAPRIIMILVHLGAPY-----QNRWLVIIMEDIANMLALNTAINPFLY 285
DB 243 VMIITPLPLFTVRYNLITLISVQDFLTHCEQSRL-PLAVQVTEVIAVTHCCNVPVY 301
QY 286 CPISRRPR 293
DB 302 AFVGERFR 309

RESULT 6
JE0296
Chyrotropin releasing hormone receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
C:Accession: JE0296
R:Itadani, H.; Nakamura, T.; Itoh, J.; Iwasa, H.; Kanatani, A.; Borkowski, J.; Ihara, Y.
Biochem. Biophys. Res. Commun. 250, 66-71, 1998
A:Title: Cloning and characterization of a new subtype of chytrotropin-releasing hormone
A:Reference number: JE0296; MUID:98407892; PMID:9735333
A:Accession: JE0296
A:Molecule type: mRNA
A:Residues: 1-352 <IT>
A:Cross-references: UNIPROT:O88820; DDBJ:AB015645; NID:g3660553; PIDD:BA33437.1; PID:g3
C:Superfamily: adenosine receptor A1
P:26-46/Domain: transmembrane #status predicted <TM1>
P:38-80/Domain: transmembrane #status predicted <TM2>
P:97-118/Domain: transmembrane #status predicted <TM3>
P:142-165/Domain: transmembrane #status predicted <TM4>
P:188-209/Domain: transmembrane #status predicted <TM5>
P:252-273/Domain: transmembrane #status predicted <TM6>
P:282-304/Domain: transmembrane #status predicted <TM7>

Query Match 12.1%; Score 226; DB 2; Length 352;
Best Local Similarity 23.5%; Pred. No. 2.1e-11;
Matches 76; Conservative 61; Mismatches 104; Indels 82; Gaps 11;

QY 31 VYSLILC-IGLPANILTVIISQLVARROKSSYN-YLLAALADILVFFIVDFLE 88
DB 26 VFLVLLVCTGLIGVAMVILV--LTSRDHPTNCVLISLADLVL----- 72
QY 89 DFLNMQMPQVDPKI-----IEVLEFSIHISIVTVELTDRIYANGCHPK 135
DB 73 ---LAAGLPVSVSLVGHMTYGRAGCLGITYOYLGNVSSFTILAFVETRYAICHPIR 129
QY 136 YHTVSPARTKAVIVSVITTCPLTSPYVWPMIWTEDYISTSVHVLIMHCFVTVLP 195
DB 130 AQVCTVABAKRIIAGIWTGVSILYCLMFPVLVDANRDNQRLCEGKVGRLPYLLD 189
QY 196 CSIFEL--LNSIIVKL-----RKSNFRLR 219
DB 190 FAVFFGLPVLTVLVLGLGRILFQSPLEQAMQKEROHQSEAAFGNCSPAKS----- 244
QY 220 GYSTGKTAALFTTSIFATLMAP-RIIMILYHLYGAPIQNRMLVHMSDILNMLALNT 278
DB 245 --SRKATMLANVLLFVLMTPRYTLYLANSFVAQPLDPLWVL----FCRTCYTNS 298
QY 279 AINFLYCFISKRFTMAATLK 301
DB 299 AVNPVVSIMSOKFR---AAFLK 318

RESULT 7
JC7319
probable allatostatin receptor-2 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004
C:Accession: JC7319
R:Leent, C.; Williamson, M.; Grimelikhuijzen, C.J.P.
Biochem. Biophys. Res. Commun. 273, 571-577, 2000
A:Title: Molecular cloning and genomic organization of a second probable allatostatin re
A:Reference number: JC7319
A:Accession: JC7319
A:Molecule type: mRNA
A:Residues: 1-357 <LEN>
A:Cross-references: UNIPROT:Q9NBC8; GB:AF25352
C:Comment: This receptor, belonging to the insect allatostatin neuropeptide family, whic
A:Genetics: A transmembrane glycoprotein.
C:Genetics: A:Gene: dar-2
A:Map position: right arm of chromosome 3, 98D-E
A:Introns: 180/3; 273/3; 326/3

C:Superfamily: endothelin receptor B
C:Keywords: extracellular protein; glycoprotein; neuropeptide; transmembrane protein

Query Match 12.0%; Score 225; DB 2; Length 357;
Best Local Similarity 24.9%; Pred. No. 2.5e-11;
Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

QY 8 LANSLSWSPQSGACGFVPPVYSSLLCLGLPANILTVIISQLVARROKSSYNLL 67
DB 31 LAINGLTLPWI-----VGF-----FFGVATGFGNLLVILV--VFNNRRSTNLM 77
QY 68 A-LAADIILVFFIVDFLEDFIL-----NMQMPQVDKIIEVLEFSIHRS 115
DB 78 VNLAAIDL--MVLICITPTATDVMYVYPRFCRSQY-----LIVTAFASTYI- 128
QY 116 IWITVPLTRRYAACHPKATVSPARTKAVIVSVITTCPLTSPY----- 163
DB 129 ---LVMSIDRLAVHPIRSRMRTENTILAIYTLVIVLVSPVAFTHDVPVDDA 185
QY 164 -----YVPMIWTEDYISTSVHVLIMHCFVTVLPVPSIFILNSIIVKLRK----- 213
DB 186 KNITVGMCTFTTNDPLGPRYQVTFPI---SSYLLPLMISGLYMRIMRLMRQGTGR 242
QY 214 -SNFLRGSYTGKTAALFTTSIFATLMAP-RIIMILYHLYGAPIQNRMLVHMSDI-A 270
DB 243 MSKESQGRK--RVTRLVAVVVIAPASLMLPVQLTLILKSL--DVIENTLTKVIVQYA 298
QY 271 NMLALNTAINFLYCFISKRFTMAATLKAFK---COKOPVOFTYN 316
DB 299 QTLAVSSCINPLVAFISENFR-----KAFYAVVCSNR-YQNTYS 339

RESULT 8
UN0605
somatostatin receptor 4 - human

C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: UN0605; UN0762; A47457
R:Xu, Y.; Song, J.; Bruno, J.F.; Bejelowitz, M.
Biochem. Biophys. Res. Commun. 193, 648-652, 1993
A:Title: Molecular cloning and sequencing of a human somatostatin receptor, hSSTR4.
A:Reference number: UN0605; MUID:93290656; PMID:8512564
A:Accession: UN0605
A:Molecule type: DNA
A:Residues: 1-388 <XU>
A:Cross-references: UNIPROT:P31391; GB:L14856; NID:9292499; PIDD:AAA36623.1; PID:9292500
R:Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; L
Biochem. Biophys. Res. Commun. 195, 844-852, 1993
A:Title: Cloning, functional expression and pharmacological characterization of a fourth
A:Reference number: UN0762; MUID:93248256; PMID:8483934
A:Accession: UN0762
A:Molecule type: DNA
A:Residues: 1-388 <YAM>
A:Cross-references: GB:D16826; NID:9693907; PIDD:BA04106.1; PID:9693908
R:Roher, L.; Raulf, F.; Bruns, C.; Buettner, R.; Hofstaedter, F.; Schule, R.
Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993
A:Title: Cloning and characterization of a fourth human somatostatin receptor.
A:Reference number: A47457; MUID:93248256; PMID:8483934
A:Accession: A47457
A:Molecule type: DNA
A:Residues: 1-82, 'T', 84-364, 'K', 366-388 <ROH>
A:Cross-references: GB:L07893; NID:g307429; PIDD:AAA60565.1; PID:g307430
A:Note: sequence extracted from NCBI backbone (NCBI:130856, NCBI:P130858)
C:Comment: This protein mediates the diverse actions of the tetradecapide somatostatin.
A:Genetics: A:Gene: SSTR4
A:Cross-references: GDB:202662; OMIM:182454
A:Map position: 20p11.2-20p11.2
A:Introns: #status absent
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; pho
P:47-73/Domain: transmembrane #status predicted <TM1>
P:84-109/Domain: transmembrane #status predicted <TM2>

F:121-142/Domain: transmembrane #status predicted <TM3>
F:162-184/Domain: transmembrane #status predicted <TM4>
F:208-238/Domain: transmembrane #status predicted <TM5>
F:257-284/Domain: transmembrane #status predicted <TM6>
F:291-314/Domain: transmembrane #status predicted <TM7>
F:2/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:119-198/Diulfide bonds: #status predicted
F:161,253/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted
F:327/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 11.7%; Score 219.5; DB 2; Length 388;
Best Local Similarity 22.0%; Pred. No. 8e-11;
Matches 72; Conservative 75; Mismatches 130; Indels 51; Gaps 11;

QY 19 PGSAGLGFVYV-VYSLICLGLPANILTVIISQLVARQKSSNYLLAADIIVL 77
DB 39 PGARAGAWAIOICIAVLCVGLVGNALVIFVILR-YAKMKATATNYLLNLVADEL-- 95
QY 78 FFIIVDFLEDFIL-NMQMPQYDKIEVLEFSSITSIWITVPLTIDRYIAVCHPKY 136
DB 96 -FHLSPFVAASSALNHMPGSLVLCRAVLSVDGLNMTSVFCLTVLSVDRYVAVHPLRA 154
QY 137 HTVSPARTKRVIVSVYITCFLTSIPY-----YWMPN-ITWEDYIST 177
DB 155 ATRRRSVAKLIMLGWMLASLVTLPIAFADTRPARGQAVACNLQMPHMAVAVV-- 212
QY 178 SVHHLIWHCFVY-VLVPCSIFFILNSIIVYKLR--KSNFRLNGYSTGKTALIFTI 233
DB 213 -----VTFPLGLFLPVLAIGLCYLIVGKRAVALPAGQORRSSEKITTRLVLA 264
QY 234 TSIFATLMPRIIMILYHLCGAPICQRMVHIMSDIANMLALNTAINFLYCFISKR 293
DB 265 VVVFVLCMPFPYVQLNLVVTSLD-----ATVNHVSLISLVANSCANPILYGLSDNFR 319
QY 294 -----TMAATLKAFFKCKQKOPQFY 314
DB 320 RSFRQVLCIRCCLEGGAGAEERPDY 347

RESULT 9
A:7249
brain-specific somatostatin receptor SSTR-4 - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A47249
R:Bruno, J.F.; Xu, Y.; Song, Y.; Berelowitz, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 11151-11155, 1992
A:Title: Molecular cloning and functional expression of a brain-specific somatostatin receptor
A:Reference number: A47249; PMID:93087484; PMID:1360663
A:Accession: A47249
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-384 <RRU>
A:Cross-references: UNIPROT:P30937; GB:M96544; NID:G207072; PION:AAA42180.1; PID:G207073
A:Note: sequence extracted from NCBI backbone (NCBI:119731, NCBI:119732)
C:Suprafamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 11.7%; Score 219; DB 2; Length 384;
Best Local Similarity 20.8%; Pred. No. 8.7e-11;
Matches 75; Conservative 80; Mismatches 139; Indels 66; Gaps 11;

QY 3 HTHAHAAANSSLSWSPG-----SACGLGFVYV-----YSLICLGLPAN 44
DB 2 NHPATPLGSEDPTTNGINASWAPDEEDAVRSDDGTAGMTTIOCIYALVCLVGVGN 61
QY 45 ILTVIISQLVARQKSSNYLLAADIIVLFIIVDFLEDFILNQMPQYVDKII 104
DB 62 ALVIFVILR-YAKMKATATNYLLNLVADELPMLSVPEVASAAA--LRHMPGAVLCRAV 118
QY 105 EVLEFSSITSIWITVPLTIDRYIAVCHPKYTVSPARTKRVIVSVYITCFLTSIPY- 163
DB 119 LSVDGIMFVSFCLTVLSVDRYVAVHPLRAATYTRPSVAKLINIGWMLASLVTLPIA 178

QY 164 -----YWMPN-ITWEDYISTSVHHLIWHCFVY-VLVPCSIFFILNS 204
DB 179 VFADTRPARGAEVACNLMPHMAVAVV-----IYFLLGLFLPVLAIGLCYL 228
QY 205 IIVYKLR--KSNFRLNGYSTGKTALIFTITSIPATLMAPRIIMILYHLCGAPIONRW 261
DB 229 LIVGKRAVALPAGQORRSSEKITTRLVLAIVTVVLCMPFPYVQLNLVFTSLD--- 285
QY 262 LVHIMSDIANMLALNTAINFLYCFISKR-----TMAATLKAFFKCKQKOPQFY 314
DB 286 -ATVNHVSLISLVANSCANPILYGLSDNFRSFRVLCIRCCLETTGGAEEERPDY 343

RESULT 10

JC7677
allatostatin receptor - American cockroach
C:Species: Periplaneta americana (American cockroach)
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: JC7677
R:Auerwald, L.; Blythe, N.; Gaede, G.; Kriengkamp, H.U.; Richter, D.
Biochem. Biophys. Res. Commun. 282, 904-909, 2001
A:Title: Structural, functional, and evolutionary characterization of novel members of
A:Reference number: JC7677; PMID:21250673; PMID:11352636
A:Accession: JC7677

A:Molecule type: mRNA
A:Residues: 1-423 <AUB>

A:Cross-references: UNIPROT:Q964D4; GB:AF336364
C:Comment: This receptor, for the neuropeptide, is involved in the inhibition of juvenile

C:Gene: Alstr

A:Keywords: neuropeptide; transmembrane protein
F:73-94/Domain: transmembrane #status predicted <TM1>
F:105-127/Domain: transmembrane #status predicted <TM2>
F:143-164/Domain: transmembrane #status predicted <TM3>
F:165-167/Region: active element DRP
F:187-207/Domain: transmembrane #status predicted <TM4>
F:243-264/Domain: transmembrane #status predicted <TM5>
F:290-312/Domain: transmembrane #status predicted <TM6>
F:327-350/Domain: transmembrane #status predicted <TM7>

Query Match 11.6%; Score 218; DB 2; Length 423;
Best Local Similarity 24.3%; Pred. No. 1.2e-10;
Matches 86; Conservative 74; Mismatches 124; Indels 70; Gaps 18;

QY 28 VPVVYSLICLGLPANILTVIISQLVARQKSSNYLLAADIIVLFEIVF-VDF 85
DB 71 VPLL-FGLIVGLFENALVLA--ANQQRSTNLLIIMLAVDLL--FIVCFVPE 124
QY 86 LLEDPLNM-QMPQYVDKIEVLEFSSITSIWITVPLTIDRYIAVCHPKYTVSPAR 144
DB 125 TATDYVLPMPFGDICKIVQYLVTAASVYTVLMSIDFELAVHDTSMSTERN 184
QY 145 TRKVIASVYITCFLTSIPYVWMPNITWEDYISTSVHHLIWHCF----- 188
DB 185 ALAALAVTVVILLASVPYLSHGEVTVY--SSAHTACVLEADINRPPDGKRVFOI 243
QY 189 ---FVYLVPCSIFFILNSIIVYKLRKSNFRLNGYST-----KTALIFTI 233
DB 244 IFPATSYVPLALICGLVIMLVRL-----WRGAAPGSHVASBRGKRTVRAVVV 296
QY 234 TSIFATLMP-RIMILYHLCGAPICQRMVHIMSDIANMLALNTAINFLYCFISK 292
DB 297 VAIFAWCPPIQLILVLSVDKRYETNTSV--MIQVSHVLAVMSCVPIIYAFISDH 354
QY 293 RTMAATLKAFFK-----OKOPQYFTNNS-ITSSPWIS--PANSHCIK 336
DB 355 R-----KAFKRVINGSGAQAQAPQEPY--HRASTIQQPAANGBALNNECV 399

RESULT 11
T18704
hypothetical protein B0334.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18704
R:Swindurne, J.
Submitted to the EMBL Data Library, October 1995
A:Reference number: 219009
A:Accession: T18704
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1340 <MIL>
A:Cross-references: UNIPROT:Q17478; EMBL:Z66519; NID:G1044812; PIDN:CAA91374.1; GSPDB:GN
A:Experimental source: clone B0334
C:Genetics:
A:Gene: CESP:B0334.6
A:Map position: 2
A:Introns: 136/1, 160/3, 307/1
C:Superfamily: Caenorhabditis elegans hypothetical protein B0334.6

Query Match 11.6%; Score 217.5; DB 2; Length 340;
Best Local Similarity 23.1%; Pred. No. 1e-10;
Matches 84; Conservative 55; Mismatches 128; Indels 97; Gaps 13;

QY 5 HAHLANSLSWSPGACGLFVPPVYSSLLCLGLPANILTVIILSQLVARQKSSYN 64
DB 31 HEHDEQISISIVWMS-----NVAVLPIY-----ALIGLACILMAVLTINKTARRIPSMN 80
QY 65 YLALAAADILVLFIVFDLFLEDFILNMQMPQVDKILEVFEFSIHSTMTITPLTI 124
DB 81 LILALAVCSLFLIFAT-----LDVTPILSIS-----LAFSTSNHFYSRIVLVI 125
QY 125 -----DRIYAVCHPLKXHTVSYPARTRKIVSVYITCFLTSPYWMFN 168
DB 126 RFLASTFYKSSRNIPKSKFLANQILE-----KIRVYRNHYR----- 164
QY 169 IWMEDYISGVHVVLIWHCFYVLYVPCSTFIILNSIIYVKLRKSNFPLRGYS----- 222
DB 165 --TMDVVSIFAFVVL-----PIGLYMSNRILFTLRRVVDSDSRKKEETKLS 211
QY 223 -----TGKTYAILFTTISIPATLMAPRI-IMILYHLYGAPIONRMVLVHIMSDIA 270
DB 212 GLIQHDAHNNRTIRANAMLFVAVFVMLFCVCGQAPARILFDWYG-QYHPKALILYC--LS 268
QY 271 NMLALINTAINFLYFCIFKREFTMAATLKAF-KQOKQVPQFYTNHNSITSSPMIS 329
DB 269 QQIVFLNASLNFCLYCVSVKRYRTLMKQTLKFKLHLEGVDPHFQNLKQYSSSAHVTS 328
QY 330 ANSH 333
DB 329 LEDH 332

RESULT 12
JC7209
Galanin receptor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C:Accession: JC7209
R:Lenz, C.; Sondergaard, L.; Grimmelikhuijsen, C.J.P.
Biochem. Biophys. Res. Commun. 269, 91-96, 2000
A:Title: Molecular cloning and genomic organization of a novel receptor from Drosophila
A:Reference number: JC7209; MUID:20160456; PMID:10694483
A:Accession: JC7209
A:Molecule type: mRNA
A:Residues: 1-394 <LEN>
A:Cross-references: UNIPROT:Q9U721; GB:AF220216
C:Comment: This receptor is a G-protein-coupled receptor and a transmembrane protein.
C:Genetics:
A:Cross-references: FlyBase:FBgn0028961
A:Map position: X distal end
A:Introns: 98/1, 161/2, 193/2, 243/3, 283/2, 308/3, 360/3
C:Superfamily: Vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

Query Match 11.6%; Score 217.5; DB 2; Length 394;
Best Local Similarity 25.9%; Pred. No. 1.2e-10;
Matches 99; Conservative 68; Mismatches 136; Indels 79; Gaps 22;

QY 2 EHT-NAHLAANSLSWSPGACG--LGFVPPVYSSLLCLGLPANILTVIILSQLVAR 58
DB 47 EHTDSDHNAANDSMEDASVALERIVSTIVEFGITGFAGLGNGLVILV--VXANQ 103
QY 59 Q-KSSYNVLLA-LAADIIVLEFIVF-VDFLEDFILNMQMPQV-----DKIEVLEFS 110
DB 104 QMRSTTNILLIINLAVSDIL--FVIFCVFPTATDVVL-----PEWFGVWVGCFVQYMLIV 156
QY 111 SIHTSIWITVPTIDRIYAVCHPLKXHTVSYPARTRKIVSVYITCFLTSP----- 162
DB 157 TCHCSVYTLVLMSPDFRLAVHVPVYMSLRERNATLIMCAMWITVTALIPVLSHSVR 216
QY 163 -YYWPNITWEDYISTSVHVLIV-----IHCF-TVYLVPCSTFIPLNSIIYVCL----- 210
DB 217 IYQHGNGACTACVFTEEE---IMSLVGQVSPFLSSVVAPELTICFLYMGMLARLWISA 273.
QY 211 -----RKSNFPLRGYSTGKTTAILFTTISIPATLMAP-RIIMIL--YHLYGAPION 259
DB 274 PCKKPSAESRKQKRV-----TRVVVVVLAFAICMLPIHVLVLKALNLVG-----G 321
QY 260 RFLVHIMSDIANMLALNTAINFLYCFISKREFTMAATLKAFK--CQKQVPQFYTN 316
DB 322 SHLSVITQIISHVAVYNTSCINPIIYAFLSDFR-----KAFKRVWCGSP-PLMTN 373
QY 317 HNFSTSSPWISPANSHCIKML 338
DB 374 QQVTKTTRTATNGTNSN-IEML 394

RESULT 13
JC8011
G protein-coupled neuropeptide pyrokinin-2 receptor (CG8784) - fruit fly (Drosophila mel
C:Species: Drosophila melanogaster
C:Date: 04-Apr-2004 #sequence_revision 04-Apr-2004 #text_change 04-Apr-2004
C:Accession: JC8011
R:RoosenKilde, C.; Gazzamali, G.; Williamson, M.; Hauser, F.; Sondergaard, L.; Delotto, R
Biochem. Biophys. Res. Commun. 309, 485-494, 2003
A:Title: Molecular cloning, functional expression, and gene silencing of two Drosophila
A:Reference number: JC8011; PMID: 12951076
A:Accession: JC8011
A:Molecule type: mRNA
A:Residues: 1-658 <ROS>
A:Cross-references: GB:AY277898
C:Comment: This receptor is a G protein-coupled receptor and a transmembrane protein as
evolution, diapause, feeding, and behavior.
C:Genetics:
A:Gene: cg8784
A:Introns: 160/1, 215/2, 259/3, 326/1, 400/3
C:Keywords: G protein-coupled receptor; neuropeptide pyrokinin; transmembrane protein

Query Match 11.6%; Score 217.5; DB 2; Length 658;
Best Local Similarity 25.0%; Pred. No. 2e-10;
Matches 82; Conservative 63; Mismatches 132; Indels 51; Gaps 14;

QY 31 VYISLLCLGLPANILTVIILSQLVARQKSSYNVLLAADIIVLFIVFDLFLEDF 90
DB 112 VCVALIFVAGVGLNLTICIVISRNPMHTATNF-YLFNLAVSDILV-----VSGIQE- 164
QY 91 ILNMQMPQV--PDKII---EVLFEFSIHSTMTITVLTIDRIYAVCHPLKXHTVSIPAR 144
DB 165 LYNLMYEDMPEFTDAMCISGVLSKMANNTVLTITAFVETRYATLCHPFRQHTMSLRS 224
QY 145 TRKIVSVYITCFLTSPYTWMPNI--WTEDYISTSVHVLIVHCFIV-----YLVPCS 197
DB 225 AIKRTFAIMWLAFLALPQAMQSVVYQNGISCTMEND--FAHVAIVAGGFIFFGSPMT 282
QY 198 IFTLNSIIYVKLRKSNFPL-----RGYST-GKTYAILFTTISIPATLMAPR 244
DB 283 AICVLVYLIGVKLRS---RLQSLPRTEDANRGLNAQGRVIRMLVAVAVAFFLCWAPF 339

QY 245 IIMILYHLGAPQIKRNL-----VHIMSDIANMLANTAINFLYCPISKRRPTMA 236
 DB 340 HAQRLAAVGLGININIGISDAFNDYPRILIDYTSGLYLSTCINPLLYIMSHKREAR 399
 QY 297 AATLKAFKCKQK----QPVQFYTNHNS 320
 DB 400 KITLTROFGIARHNNHQQSQNH-QHNYS 426

RESULT 14

JC2421
 opioioid receptor homolog, MOR-C - mouse

C/Species: Mus musculus (house mouse)
 C/Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
 C/Accession: JC2421; 149122
 R/Mishl, M.; Takeshima, H.; Mori, M.; Nakagawara, K.; Takeuchi, T.
 Biochem. Biophys. Res. Commun. 205, 1353-1357, 1994
 A/Title: Structure and chromosomal mapping of genes for the mouse kappa-opioid receptor
 A/Reference number: JC2434; MIMD:95100967; PMID:7802669
 A/Accession: JC2421
 A/Molecule type: mRNA
 A/Residues: 1-367 <NIS>
 A/Cross-references: UNIPROT:P35377; DDBJ:D31663
 R/Halford, W.P.; Gebhardt, B.M.; Carr, D.J.
 J. Neuroimmunol. 59, 91-101, 1995
 A/Title: Functional role and sequence analysis of a lymphocyte orphan opioioid receptor.
 A/Reference number: 149122; MIMD:95318231; PMID:7797625
 A/Accession: 149122
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: mRNA
 A/Residues: 1-357 <RES>
 A/Cross-references: EMBL:U14165; NID:G540092; PID:AAA87899.1; PID:G540093
 C/Genetics:
 A/Gene: MOR-C; OOR
 A/Map position: 2H2-4
 A/Introns: 75/2; 194/1
 C/Superfamily: vertebrate rhodopsin
 C/Keywords: receptor

Query Match 11.6%; Score 217; DB 2; Length 367;
 Best Local Similarity 22.8%; Pred. No. 1.2e-10;
 Matches 74; Conservative 65; Mismatches 117; Indels 68; Gaps 13;
 QY 7 HLLAANGSLSWSPGSAAGLGF-VPVYYSGLLCL-GLPANIILTVIILSOLVARROKSSYN 64
 DB 32 HLLANSHSAFLP-----LGLKVTIVGLYLAVCIGGLGNCIMVYI--LRHTKMTATN 84
 QY 65 -YLLAALADIIIVLPPIV--VDFLEDFILNQMPQVPDKIIEVLEFSSIHSTIWTVP 121
 DB 85 IYFNLALADTLVLTLPFGTDILGFWPFGNAL-----CKTVIADYMMFTSTFTLTA 140
 QY 122 LITDRYIANCHPKYHTVSPARTKVIYSVITCTLSIPY----- 163
 DB 141 MSYDRYVALCHPRALDVRTSSKAQAVNVAIMALASVGVPAIMGSAQVEDEIECLVE 200
 QY 164 -----YMPNITWEDYISTSVHVLIMHCFVTVLPCISIFPLINSIIVYKLR----- 211
 DB 201 IPAPQDYWGP-----VFAICIFLPSF--IIPVLIISVCISLMIRRLRGVRLLS 246
 QY 212 --RKSNFLRGYSTGKTTAILFTTISIFATLMAPIIMILYHLYGAPIQNRMLVHIMSDI 269
 DB 247 GSREKDRNLR-----RITRLVLVVAVFPVGCWTPVQVFLVQGLGVQPSGSETAVAILR-P 300
 QY 270 ANMLALNTAINFLYCFISKRR 293
 DB 301 CTALGVNSCLNPILVAFIDENFK 324

RESULT 15
 149022
 kappa opioioid receptor 3 - mouse
 C/Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C/Accession: 149022
 R/Pan, Y.X.; Cheng, J.; Xu, J.; Rosei, G.; Jacobson, E.; Ryan-Moro, J.; Brooks, A.L.; D
 Mol. Pharmacol. 47, 1180-1188, 1995
 A/Title: Cloning and functional characterization through antisense mapping of a kappa 3
 A/Reference number: 149022; MIMD:95327076; PMID:7603458
 A/Accession: 149022
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: mRNA
 A/Residues: 1-367 <RES>
 A/Cross-references: UNIPROT:P35377; EMBL:U09421; NID:G551484; PID:AAA81333.1; PID:G551
 C/Superfamily: vertebrate rhodopsin

Query Match 11.6%; Score 217; DB 2; Length 367;
 Best Local Similarity 22.8%; Pred. No. 1.2e-10;
 Matches 74; Conservative 65; Mismatches 117; Indels 68; Gaps 13;
 QY 7 HLLAANGSLSWSPGSAAGLGF-VPVYYSGLLCL-GLPANIILTVIILSOLVARROKSSYN 64
 DB 32 HLLANSHSAFLP-----LGLKVTIVGLYLAVCIGGLGNCIMVYI--LRHTKMTATN 84
 QY 65 -YLLAALADIIIVLPPIV--VDFLEDFILNQMPQVPDKIIEVLEFSSIHSTIWTVP 121
 DB 85 IYFNLALADTLVLTLPFGTDILGFWPFGNAL-----CKTVIADYMMFTSTFTLTA 140
 QY 122 LITDRYIANCHPKYHTVSPARTKVIYSVITCTLSIPY----- 163
 DB 141 MSYDRYVALCHPRALDVRTSSKAQAVNVAIMALASVGVPAIMGSAQVEDEIECLVE 200
 QY 164 -----YMPNITWEDYISTSVHVLIMHCFVTVLPCISIFPLINSIIVYKLR----- 211
 DB 201 IPAPQDYWGP-----VFAICIFLPSF--IIPVLIISVCISLMIRRLRGVRLLS 246
 QY 212 --RKSNFLRGYSTGKTTAILFTTISIFATLMAPIIMILYHLYGAPIQNRMLVHIMSDI 269
 DB 247 GSREKDRNLR-----RITRLVLVVAVFPVGCWTPVQVFLVQGLGVQPSGSETAVAILR-F 300
 QY 270 ANMLALNTAINFLYCFISKRR 293
 DB 301 CTALGVNSCLNPILVAFIDENFK 324

Search completed: December 14, 2004, 20:25:45
 Job time : 16.4373 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2004, 20:28:19 ; Search time 2586 Seconds
(without alignments)
4974.178 Million cell updates/sec

Title: US-09-995-225B-16

Perfect score: 1872
Sequence: 1 MEHTAHILANSLSMWSRC.....CIKMLVVOYDKNGKPIKVP 353

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+pn.model -DEV=xlh
-O=/cgn2.1/USPTO.epoc1/US09995225/runat.14122004.110733.18040/epc.query.fasta_1.519
-DB=EST -QFMT=fastlap -SUFFIX=est -MIMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITs=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09995225@CGN_1_1.2607@runat.14122004.110733.18040 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WMAP TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST+
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	925	49.4	803	9	CC905917 t025011ba
2	606	32.4	806	9	CNS02284 AL178069 Tetracodon
3	509	27.2	824	9	CL137737 ISB1-110A
4	506	27.0	362	8	AZ513631 1M0359L07
5	453.5	24.2	947	9	CNS02BVQ AL190367 Tetracodon
6	444	23.7	867	9	CL137848 ISB1-110C
7	433.5	23.2	813	8	BZ169352 CH230-259
8	426	22.8	898	9	CNS02ENP AL193966 Tetracodon
9	414.5	22.1	515	1	AL921815 AL921815

C 10	412	22.0	448	8	AQ225693 HS 2009, B
11	380	20.3	1269	5	CL649540 CH213-237
12	359	19.2	708	5	BU610327 UI-M-DJ2-
C 13	314	16.8	588	8	AQ344035 RRC11-12
14	309	16.5	603	8	AQ351433 RRC11-11
15	249	13.3	1107	9	CNS04L7F AL295764 Tetracodon
16	241	12.9	1197	9	AY398777 Nus muscu
17	239.5	12.8	1197	9	AY398776 Pan trogl
18	238.5	12.7	1197	9	AY398775
19	231.5	12.4	1086	9	AY399291
20	231.5	12.4	2583	3	CR609171 full-leng
21	218	11.6	1146	9	AY414479 Nus muscu
22	218	11.6	2785	3	AK028281 Nus muscu
23	217	11.6	2919	3	AK038620 Nus muscu
24	217	11.6	2959	3	AK079529 Nus muscu
25	216.5	11.6	2432	3	AK051189 Nus muscu
26	216.5	11.6	2708	3	AK076275 Nus muscu
27	215.5	11.5	1086	9	AY399293 Nus muscu
28	215.5	11.5	2118	3	AK031109 Nus muscu
29	215.5	11.5	2292	3	AK036597 Nus muscu
30	215.5	11.5	2432	3	AK036690 Nus muscu
31	213	11.4	863	5	BP161654 BP161654
32	213	11.4	1511	3	CR601301 full-leng
33	213	11.4	1603	3	CR614199 full-leng
34	213	11.4	1612	3	CR594588 full-leng
35	213	11.4	1619	3	CR610268 full-leng
36	213	11.4	1621	3	CR614594 full-leng
37	213	11.4	1626	3	CR598773 full-leng
38	213	11.4	1629	3	CR596547 full-leng
39	213	11.4	1629	3	CR623838 full-leng
40	213	11.4	1640	3	CR619476 full-leng
41	213	11.4	1642	3	CR594428 full-leng
42	213	11.4	1650	3	CR605121 full-leng
43	213	11.4	1682	9	CR614663 full-leng
44	207.5	11.1	1662	9	AY400424 Nus muscu
45	207.5	11.1	1200	9	AY416064 Homo sapi

ALIGNMENTS

RESULT 1
CC905917/c
LOCUS
DERIVATION
CC905917 803 bp DNA linear GSS 08-AUG-2003
t025011ba.r1 TAMER Bos taurus genomic clone t025011ba, genomic
survey sequence.

ACCESSION
CC905917 GI:33524850

VERSION
CC905917.1 GI:33524850

KEYWORDS
GSS.

SOURCE
Bos taurus (cow)

ORGANISM
Bos taurus

REFERENCE
1 (bases 1 to 803)

AUTHORS
Lin, S., Najjar, F.Z., Adelson, D., Gill, C.A. and Roe, B.A.

TITLE
Bovine BAC End Sequences from Library TAMER

JOURNAL
Unpublished (2003)

COMMENT
Contact: Bruce A. Roe

Advanced Center for Genome Technology
University of Oklahoma Department of Chemistry and Biochemistry
620 Parrington Oval, Room 208, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu

Class: BAC ends
High quality sequence start: 39
High quality sequence stop: 551.

FEATURES

source
1..803
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Angus bull T A M U Shoshone Y6 11519666"
/db_xref="taxon:9913"

QY 177 ThrSerValHisHisValleuileTrrPleHisCyephenthValTyrluValProCyS 196
Db 365 AMGGCCCTGGAGCCGCTCCATCTGAGCCAMGTGACATCATCTTCTGCGCCCTGC 424
QY 197 SerIlePhePheIleleuanserIleleValTyrluValArg--ArgIysSerAsn 215
Db 425 AGCATTTCTGCTGCTCAACTTTGATCATCCMAAGCTGAGGCGGAGAGCGG 484
QY 216 PheArgLeuArgGly-----TyrSerThrGlyLysTrp 226
Db 485 CAGCGCAG--CGGGGCGAGCGCGGCMCAAGTMCACGCTCCGCGCGCTGAGAGAG 543
QY 227 ThrAlaIleleuPheThrIleThrSerIlePheAlaThrleuTrpAlaProArgIleIle 246
Db 544 ACGGCATGCTGCTGGCCATCATCTCCGCTTCTGCTGCTGCGGACGACGAGCGTG 603
QY 247 MetIleleuTyrlHisleuTyrglyAlaProIleGlnAsnArgTrpLeuValHisIleMet 266
Db 604 GTGGTCATCTACACACCTGTACGCTGATGTGCTTCAACGCGAGTGGCGGCTGCACCTGGCC 663
QY 267 SerAspIleAlaAsnMetIleuAlaLeuLeuAsnThrAlaIleAsnPheleuTyrgly 286
Db 664 TACGACCTGTCCCAACATGCTGCGCATGTCTCAACACGCGCGTAACCTTCTGCTACTCG 723
QY 287 PheIleSerIysArgPheArg 293
Db 724 TTGTCAGCAAGCGCTTCCGC 744

RESULT 3
LOCUS CL137737 824 bp DNA linear GSS 05-JUN-2004
DEFINITION ISB1-110A2.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-110A2,
genomic survey sequence.
ACCESSION CL137737
VERSION CL137737.1 GI:40631372
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 824)
REFERENCE
AUTHORS Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submls@wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 2
High quality sequence stop: 672.

FEATURES
source
1..824
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone_1lb="ISB1"
/clone_1lb="ISB1-110A2"
/note="Vector: pBelOBAC11, ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN
Alignment Scores:
Pred. No.: 2.3e-43 Length: 824
Score: 509.00 Matches: 97
Percent Similarity: 59.51% Conservative: 50
Best Local Similarity: 39.27% Mismatches: 69
Query Match: 27.19% Indels: 31
DB: 9 Gaps: 3

US-09-995-225B-16 (1-353) x CL137737 (1-824)
QY 27 PheValProValValTyrlTyrlSerleuLeuCyelGlyLeuProAlaAsnIleleu 46
Db 157 TTTCTTCTTACAGTG-----AACATGCTG 180
QY 47 ThrValIleIleleuSerGlnleuValAlaArgArgGlnLysSerSerTyrlAsnTyrlleu 66
Db 181 ACTGCTGTGACCTCTCAAGGCTGGAGCGAGAAAGAGTCCCTTACAGTACCTG 240
QY 67 LeuAlaLeuAlaAlaIlePheleuValleuPhePheIleValPheValAspPheleu 86
Db 241 TTGGCTCTTACCATCTCCACACATCTTGACCCAGATTTTCATCTTTTGGGCTTATC 300
QY 87 LeuGluAspPheIleleuAsnMetGlnMetProGlnValProAspLysIleIleGluVal 106
Db 301 CTGCAAAACAGAACTACTCCACCGCAAGGCGCCAGTACCTTAATCATGTGTCAGTATGC 360
QY 107 LeuGluPheSerSerIleHisThrSerIleTrrPleHisValProleuThrIleAspArg 126
Db 361 CTTGAGTTCTCTCTTAATCATCATCTTATTTGGGTGACGTATACAGCTGACCGGT 420
QY 127 TyrlleAlaValCyHisIleProleuTyrlHisThrValSerTyrglyProAlaArgThrArg 146
Db 421 TATGTGGCATTTATGTCATCTCCGTACATATGCTCTTCTTACCCAGAGCAACCGGT 480
QY 147 LysValIleValIleValSerValTyrlleThrCyPheleuThrSerIleProTyrlTyrrP 166
Db 481 AGAGTCATGTCTTGTCTTCTCCATCATCTTTGTGACCGGTGATCCCTTCTACTGCTG 540
QY 167 ProAsnIleTrrPheGluAspTyrlleSerThrSerValHisHisValleuIleTrrPle 186
Db 541 AGTAGTGTGGAGGATCGCGTACCCAGATGTGACCTTAATCATCTTACAGTGGACC 600
QY 187 HisCyPheThrValTyrlleuValProCySerIlePhePheIleleuAsnSerIlele 206
Db 601 CACTGCTTATATATATCTTCACTTCTCCCTGACACATTTTGTGATTAACAACCTGTCATT 660
QY 207 ValTyrlLysLeuArgArgLysSer-----AsnPheArgLeuArgGly 220
Db 661 ATCTACAGGCTGAGGAAAGAAAGAAAGAGGGGTCCAGAAATTTGCAGCTTCTTAATGGGC 720
QY 221 TyrlSerThrGlyLysThrThrAlaIleleuPheThrIleThrSerIlePheAlaThrleu 240
Db 721 AAGACCCACAGCAAAATTTCTCCGGGCGCATTAACCTCTGTTTTTTCGCGGTCTTT 780
QY 241 TrrAlaProArgIleleuMetIleleuTyrlHisleuTyrglyAlaProIleGlnAsnArg 260
Db 781 TGG-----GGCTCTCTTAAAAAAACCATT 804
QY 261 -TrpLeuValHisIleMet 266
Db 805 TTGGTCATCATTTAATT 823

RESULT 4
LOCUS AZS13631/c 362 bp DNA linear GSS 05-OCT-2000
DEFINITION IM0359L07R Mouse 10kb plasmid UGCGIM library Mus musculus genomic
clone UGCGIM0359L07 R, genomic survey sequence.
ACCESSION AZS13631
VERSION AZS13631.1 GI:10694947
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D., Weis,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb

Db		72	NACTGCGCTCCGCAGCCTCGCGGGGCGCGCGTGCACATTTCTACGCATCTT	131
Oy		270	AlaansncttleualaaleuleubantThralalaeapnepheleutyrcypheilleser	289
Db		132	GCCACAACATGTTCGCGTGCTCTCAACACGGGVTGTAACTTCTTCTACAGTTTCATCAGC	191
Oy		290	LysatrgPheaYgThMeclAlalaIathrlenuysalaphenelyCyglintyagin	309
Db		192	WAGCTTTTCGGGGGCGATCGCGCCAACGTGCTGCCGAGCGCTGCTCACATGCAGGAAGACG	251
Oy		310	ProvalGlnphetyrThrasnhtisanPheserileThrserserProtripleserPro	329
Db		252	CTGCAGCGCGTCTACGCCAGCCACACACTTTTTCATCAGAGTGTGCCCCCTGmwtcACCA	311
Oy		330	AlaanssrthacYeillelymetleuualtyrygintyrasplyasaingilyabroille	349
Db		312	GCAAACCTCCCACTGCATCAAGTGTGTGGTGTACGTAACGACAAAACGGGAAGCCCATC	371
Oy		350	lysValser	352
Db		372	TGTATTTC	380
RESULT 6				
LOCUS	CLJ37848	867 bp	DNA	linear GSS 05-JAN-2006
DEFINITION	ISB1-110C15_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-110C15,			
ACCESSION	CLJ37848			
VERSION	CLJ37848.1	GI:40631483		
KEYWORDS	GSS.			
SOURCE	Xenopus tropicalis (western clawed frog)			
ORGANISM	Xenopus tropicalis			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Silurana. 1 (bases 1 to 867) Kremetzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R. A physical map of the xenopus tropicalis genome Unpublished (2003) Contact: Richard K Wilson Genome Sequencing Center, Washington University School of Medicine Email: submissions@wustl.edu Insert Length: 75000 Std Error: 0.00 Seq primer: T7 TAATACGACTCACTATAGG Class: BAC ends High quality sequence start: 11 High quality sequence stop: 673. Location/Qualifiers			
JOURNAL				
COMMENT				
FEATURES				
source	/organism="Xenopus tropicalis"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:8364"			
	/clone="ISB1-110C15"			
	/clone_1lb="ISB1"			
	/note="Vector: pBelobAC11, ISB-1 Xenopus tropicalis BAC library Segment 1"			
ORIGIN				
Alignment Scores:				
Pred. No.:	2.07e-36	Length:	867	
Score:	444.00	Matches:	92	
Percent Similarity:	61.06%	Conservative:	46	
Best Local Similarity:	40.71%	Mismatch:	62	
Query Match:	23.72%	Indels:	26	
DB:	9	Gaps:	4	
US-09-995-225B-16 (1-353) x CLJ37848 (1-867)				
Oy	27	pheValpProvalValTyrrTyrrSerleuleuleuCybleugilyleuProAlasnlleueu	46	
Db	172	TTTTCTTCTTACAGTG-----AACATGCTG	195	

QY	47	ThrValIleIleLeuSerGlnValValAlaArgAlaGlnIleuSerSerTyraSerIleu	66
Db	196	ACTGCTGTGGACACTCTCAAGGCTGGCATTCGAGAACAAAGAGCTCTCTACAGTACCTG	255
QY	67	LeuAlaLeuAlaAlaAlaSerIleLeuValIleuPhePheIleValPheValAspPheLeu	86
Db	256	TTGGCTCTTACCATTCAGACATCTTGAGCGAATTTTTCATCATTTTTGTGGGCTTATC	315
QY	87	LeuGluAspPheIleLeuMetGlnMetProGlnValProAspIleIleGluVal	106
Db	316	CTGCAAAACAGCAATATCCACCCGAGGTGCCAGTACTTAAATCCATGGTGCTAGTGC	375
QY	107	LeuGluPheSerSerIleIleIleSerIleIleThrIleThrValProLeuThrIleAspArg	126
Db	376	CTTAGATTTCTCTTAAATCATTCATCTTAATTTGGGTGCAAGTATCATCTGACCGTACCT	435
QY	127	TyrIleAlaValCysHisProLeuValYHisThrValSerTyrProAlaArgThrArg	146
Db	436	TATGTGGCATTAATGCTATTCGGCTCAATATTCGCTCTTCTTACCAAGAGCGAACCGT	495
QY	147	LysValIleValSerValTyrIleThrCysPheLeuThrSerIleProTyrTyrTrpTrp	166
Db	496	AGAGTCAATGCTCTTGGTCTCTCTCATCGTTGTGACCGGTGACCTTCTACTGTGG	555
QY	167	ProAsnIleThrPthrIleuAspTyrIleSerThrSerValHisValIleuIleTrpIle	186
Db	556	AGTCAATGTTGGAGGAGATCCCGGATCCCGAGATGCTTAACCTTAATCTCAAGTGAC	615
QY	187	HisCysPheThrValTyrLeu--ValProCysSerIlePhePheIleuAsnSerIleI	206
Db	616	CACGCTTTTAATATATCTTCATTCCTCCGACACATATTTTTTGTGATTCCAAACTCTG	675
QY	206	LevalTyr-----LysLeuValArgIleGlySer-----Asp 216	
Db	676	GTCATTTATCTTACCAAGGCTGAGAAAAAAACGGGGTCCCAAGAAAGTTGCCATCCT	735
QY	216	heArgLeuArgGlyTyrSerThrClyLysThrThrAlaIleLeuPheThrIleThrSerI	236
Db	736	TTCTCTAAATGGGCGCAAGAA-----CCACCACTCAATTTCTTCTTGGAGACAA	769
QY	236	IePheAlaThrLeu 240	
Db	790	TTTATCCCTCCCTG 803	
RESULT 7			
LOCUS	B2169352		
DEFINITION	CH230-259E22.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone		
ACCESSION	B2169352		
VERSION	B2169352.1	GI:23810547	
SOURCE	GSS.		
ORGANISM	Rattus norvegicus (Norway rat)		
REFERENCE	1 (bases 1 to 813)		
AUTHORS	Zhao,S., Shetty,J., Shatman,S., Teegeye,G., Geer,K., Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., R1998,F., de Jong,P. and Fraser,C.M.		
TITLE	Rat BAC End Sequences from Library CHORI-230 MboI segment		
JOURNAL	Unpublished (1999)		
COMMENT	Other_GSSes: CH230-259E22.TV Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: szhao@tigr.org Clones are derived from the rat BAC library CHORI-230		

MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2009 row: D column: 16
 Class: BAC ends
 High quality sequence stop: 448.
 Location/Qualifiers

FEATURES
 source 1..448
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=2009 Col=16 Row=D"
 /sex="male"
 /clone_1lb="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pReloBAC11; BAC Clones in E-Coli DH10B"

ALIGNMENT SCORES:

Pred. No.: 2,06e-33 Length: 448
 Score: 412.00 Matches: 77
 Percent Similarity: 95.29% Conservative: 4
 Best Local Similarity: 90.59% Mismatches: 4
 Query Match: 22.01% Indels: 0
 DB: 8 Gaps: 0

US-09-995-225b-16 (1-353) x AQ225693 (1-448)

QY 269 ILEAIAAAMetLeuThraAlaIleAenPheLeuTyCysPheIle 288
 Db 435 ATAGCAAGATGATACCTGTGGAGCAGCATCACTTCTTACTGCGCTATC 376
 QY 289 SerLYAaSPheArghTherMeAlaAlaArhLeuYsAlaPhePhelYsGlnYs 308
 Db 375 AGCAAGCGGTCCGACCATGCGCGCAGCTTCAGATGCCAGAG 316
 QY 309 GlnProValGlnPheYrThraSnhIaSnPheSerIleThrSerProTyrIleSer 328
 Db 315 CACCMYTAAGTTCACCAATCACTTTCATACAGAGCCCTGGATCTCG 256
 QY 329 ProAlaAsnSerHisCysIleYsMetLeuValTyrGlnTyrAspLYsAsnGlyLYsPro 348
 Db 255 CCGCAAACTCAGCATGATCAAGTCTGTGTACAGATGACAAAAAGAAACCT 196
 QY 349 IleLYsValSerPro 353
 Db 195 ATAAAAATATCCCG 181

RESULT 11

LOCUS CL649540 1269 bp DNA linear GSS 06-JUN-2004
 DEFINITION CH213-237F13.SP6 CH213 Gaesterosteus aculeatus genomic clone
 CH213-237F13 3', genomic survey sequence.
 ACCESSION CL649540
 VERSION CL649540.1 GI:49668964
 KEYWORDS GSS.
 SOURCE Gaesterosteus aculeatus (three spined stickleback)
 ORGANISM Gaesterosteus aculeatus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Gaesterosteiformes;
 Gaesterosteidae; Gaesterosteus.
 1 (bases 1 to 1269)
 AUTHORS Kingsley, D., Grimwood, J., Dickson, M., Schmutz, J. and Myers, R.M.
 TITLE Expressed sequence tags from Gaesterosteus aculeatus
 JOURNAL Unpublished (2004)

COMMENT

Contact: Grimwood, Jane
 Stanford Human Genome Center
 Stanford University School of Medicine
 975 S California Avenue, Palo Alto, CA 94304, USA
 Tel: 650 320 5917
 Fax: 650 320 5801
 Email: jane@hgsc.stanford.edu
 Plate: 237
 Class: BAC ends
 High quality sequence start: 19
 High quality sequence stop: 869.
 Location/Qualifiers

FEATURES
 source 1..1269
 /organism="Gaesterosteus aculeatus"
 /mol_type="genomic DNA"
 /strain="Salmon River"
 /db_xref="taxon:69293"
 /clone="CH213-237F13"
 /sex="Mixed"
 /cell_type="Blood"
 /clone_1lb="CH213"
 /note="Vector: pPARBAC2.1; Site: 1; EcoRI; The sequence of the clone was established as a mapping and sequencing collaboration at the Stanford Genome Evolution Center, funded by the NIH Centers of Excellence in Genomic Science (CEGS) initiative (<http://cegs.stanford.edu>). The clone was isolated from the BAC library CHORI-213 built by Peter deLong in collaboration with the Stanford Genome Evolution Center (<http://www.chori.org/bacpac/>). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm)."

ORIGIN

ALIGNMENT SCORES:

Pred. No.: 2,33e-29 Length: 1269
 Score: 380.00 Matches: 72
 Percent Similarity: 82.24% Conservative: 16
 Best Local Similarity: 67.29% Mismatches: 16
 Query Match: 20.30% Indels: 4
 DB: 9 Gaps: 1

US-09-995-225b-16 (1-353) x CL649540 (1-1269)

QY 246 ILEMeTlleLeuTyHisLeuTyrgLYaLAPollegIaAsnArgrTyrLeuValHisIle 265
 Db 11 GTTTACTCGCGCGCGCTTCACGGGGCGCGG-CGA-----CTGCTCACGTG 60
 QY 266 MetSerAspIleAlaAsnMetLeuAlaLeuAenThraAlaIleAenPhePheLeuTy 285
 Db 61 CTCACGACCTGGCCATTATGCTCGCGTTCACACCGGGGTCACACTTCTTCTTAC 120
 QY 286 CysPheIleSerLYaSPheArghTherMeAlaAlaArhLeuYsAlaPhePhelYs 305
 Db 121 TGCCTCATCAGCAAGCGTTCCGCGCATGCGCGCAACGTCCTCGCGCCCTGCTCAC 180
 QY 306 CysGlnLYsGlnProValGlnPheYrThraSnhIaSnPheSerIleThrSerPro 325
 Db 181 TCCCGAAGACGCGCCCGCTTCTACGCCAGCAACTTTCATCAGCAAGAGCCCG 240
 QY 326 TrrIleSerProAlaAsnSerHisCysIleYsMetLeuValTyrGlnTyrAspLYsAsn 345
 Db 241 TGGATCTACCGGCAACTCCCACTGCAATTAGATGCTGTGTACAGATGACAAAAAT 300
 QY 346 GlyLYsProlleYsValSer 352
 Db 301 GGGAAACCGTGTATTTCC 321

RESULT 12

LOCUS BUE10327 708 bp mRNA linear EST 20-FEB-2003
 DEFINITION UI-M-DJ2-bw1-f-07-0-UI.r1 NIH_BMAP DJ2 Mus musculus cDNA clone
 UI-M-DJ2-bw1-f-07-0-UI 5', mRNA sequence.
 ACCESSION BUE10327

```

VERSION      BU610327.1  GI:23276542
KEYWORDS     EST.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
REFERENCE    Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        1 (bases 1 to 708)
JOURNAL      Ronaldo, M.F., Lennon, G. and Soares, M.B.
MEDLINE      Normalization and subtraction: two approaches to facilitate gene
PUBMED       discovery
COMMENT      Genome Res. 6 (9), 791-806 (1996)
              97044477
              8889548
              Contact: Chin, H
              National Institute of Mental Health
              6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
              20892-9643, USA
              Tel: 301 443 1706
              Fax: 301 443 9890
              Email: mestr@mail.nih.gov
              Tissue Procurement: Dr. Robin Davison
              cDNA library preparation: Dr. M. Bento Soares, University of Iowa
              cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Researchers may obtain clones from Research
              Genetics (www.reagen.com).
              Seq primer: M13 REVERSE.
              Location/Qualifiers
                1..708
                  /organism="Mus musculus"
                  /mol_type="cDNA"
                  /strain="C57BL/6"
                  /db_xref="taxon:10090"
                  /clone="U1-M-DJ2-bw1-f-07-0-UI"
                  /cissue_type="subfornical organ and postrema"
                  /dev_stage="Adult"
                  /lab_host="DH10b (Life Technologies) (11 phage resistant)"
                  /clone_1lb="NH_BMAP_DJ2"
                  /note="Organs: Brain; Vector: pT7T3-Pac (Pharmacia) with a
                  modified polylinker; Site 1: EcoR I; Site 2: Not I;
                  U1-M-DJ2 is a subcloned cDNA library constructed
                  according to Bonaldo, Lennon and Soares, Genome Research,
                  6:791-806, 1996. First strand cDNA synthesis was primed
                  with an oligo-dT primer containing a Not I site. Double
                  stranded cDNA was ligated to an EcoR I adaptor, digested
                  with Not I, and cloned directionally into pT7T3-Pac
                  vector. The oligonucleotide used to prime the synthesis of
                  first-strand cDNA contains a library tag sequence that is
                  located between the Not I site and the (dT)18 tail. The
                  sequence tag for this library is: GCTACATGAT, subfornical
                  organ and area postrema."
ORIGIN
Alignment Scores:
Pred. No.:      1,75e-27      Length:      708
Score:          359.00        Matches:      66
Percent Similarity: 98.53%    Conservative: 1
Best Local Similarity: 97.06%  Mismatches: 0
Query Match:    19.18%       Indels:      0
DB:             5            Gaps:         0
US-09-995-225B-16 (1-353) x BU610327 (1-708)
QY      286  CysPhelIeserLyarGphearGThrMetAlaAlaAlaThrLeuYsAlaPhePhelys 305
      |||
Db      3   TECTTCATAGCAAGGCTTCCTCCATCAATGCGAGCTGCCACATCAAGGCTTGTTCAG 62
QY      306  CysGlnLyeglnProYalGlnPheYrThrAenHiaenPheSerIleThrSerSerPro 325
      |||
Db      63  TGTCAAGACAGAGCTGTACAGTCTATACCAACCATTAATTCCATTAACAAGTATGCC 122
QY      326  TPpIleSerProIlaenSerHisCysIleYemMerLeuValYrGlnYrAapYsAen 345
      |||

```

```

Db      123  TGGATCTCACCAAGAACTCACATGATCAAGATGCTGTGTACAGTATGACAAACAT 182
QY      346  GilyYpProIleYsValSerPro 353
      |||
Db      183  GGAAAGCTATTAAGATATCCCCG 206
RESULT 13
LOCUS      AO344035/c
DEFINITION RPII11-124U23-TV RPI1-11 Homo sapiens genomic clone RPI1-11-124U23,
ACCESSION  AO344035
VERSION     AO344035
KEYWORDS    AO344035.1 GI:4168931
SOURCE      GSS.
ORGANISM    Homo sapiens (human)
            Homo sapiens
            Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
AUTHORS     Venter, J.C.
TITLE       Use of BAC End Sequences from Library RPI1-11 for Sequence-Ready
JOURNAL     Map Building
COMMENT     Unpublished (1997)
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPI1-11. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@jng.med.bufileo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.bufileo.edu/ordering) or from
            Research Genetics (info@reagen.com). BAC end search page:
            http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
            Location/Qualifiers
              1..588
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="GDB:7547470"
                /db_xref="taxon:9606"
                /clone="RPI1-11-124U23"
                /sex="Male"
                /cell_type="Lymphocytes"
                /clone_1lb="RPI1-11"
                /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                RPI11 Human Male BAC Library"
ORIGIN
Alignment Scores:
Pred. No.:      8,31e-23      Length:      588
Score:          314.00        Matches:      66
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    16.77%       Indels:      0
DB:             8            Gaps:         0
US-09-995-225B-16 (1-353) x AO344035 (1-588)
QY      43  AlaAsnIleLeuThrValIleIleuSerGlnLeuValAlaArgGlnYsSerSer 62
      |||
Db      199  GCAAAATATTTAGACAGTATATATCTCTCCAGCTGTGTGCAAGAAGAGTCTTCC 140
QY      63  TyrAenYrLeuLeuAlaLeuAlaAlaAlaApIleLeuValLeuPhePheIleValPhe 82
      |||
Db      139  TACAATATCTCTTGGCACTCCCTGCGACATCTTGCTCTTTTCAATAGTATT 80
QY      83  ValAepPheLeuGlnAapPheIleLeuAenMeGlnMetProGlnValProAapYs 102
      |||

```

Db	79	GTGCACTTCCTGTGGAGAATTTCATCTTGAACATGACAGATCCCTCAGSTCCCCGACAG	20
Oy	103	le eg u v a le u g l u 108	
Db	19	ATCATAGAAAGTGCTCGAA	2
RESULT 14			
AQ351433/c			
LOCUS	AQ351433	603 bp	DNA
DEFINITION	RPCT11-113113.TV RPCT-11 Homo sapiens genomic clone RPCT-11-113113		linear GSS 07-MAY-1991
ACCESSION	AQ351433		
VERSION	AQ351433.1		
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 603)		
JOURNAL	Zhao,S., Adams,M.D., Nierman,W., Malek,T., de Jong,P. and		
COMMENT	Venter,J.C.		
	Use of BAC End Sequences from Library RPCT-11 for Sequence-Ready		
	Map Building		
	Unpublished (1997)		
	Contact: Shaying Zhao, William Nierman, Mark Adams		
	Department of Eukaryotic Genomics		
	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850		
	Tel: 301 838 0200		
	Fax: 301 838 0208		
	Email: hbe@tigr.org		
	Clones are derived from the human BAC library RPCT-11. For BAC		
	library availability, please contact Pieter de Jong		
	(pieter@dejong.med.buffalo.edu). Clones may be purchased from		
	BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from		
	Research Genetics (info@resgen.com). BAC end search page:		
	http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html		
	Seq primer: SP6		
	Class: BAC ends.		
FEATURES			
source	Location/Qualifiers		
	1..603		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="GDB:7543212"		
	/db_xref="ECSL:7543212"		
	/clone="RPCT-11-113113"		
	/sex="Male"		
	/cell_type="Lymphocytes"		
	/clone_11b="RPCT-11"		
	/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;		
	RPCT11 Human Male BAC Library"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.94e-22	Length:	603
Score:	309.00	Matches:	65
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	16.51%	Indels:	0
DB:	g	Gaps:	0
US-09-995-225B-16 (1-353) x AQ351433 (1-603)			
Oy	43	AlaAenleleuThrrvAllelleleuSerGlnleuValAlaArGArgGlnlySerSer	62
Db	195	GCAAAATCTTGACAGTATCATCTCTCCACGCTGTGGCAAGAAACAGAGTCTCC	136
Oy	63	TyAenTyrlleuAlaAlaAlaAlaAlaAspilleuValleuPhePheillevalPhe	82
Db	135	TACACATATCTCTTGACATCGCTGCTCCGACATCTTGCTCTTTTTCATATATGTTT	76
Oy	83	ValAspPheleuLeuGlnAspPheilleuAsnMetGlnMetProGlnValProAspLys	102

Db		75	GTGACCTTCCTTGGAAGATTTCATCTTGAACATGCAGATGCTCAGGTCGCCAAGA	16	
Oy	103	lelelgiuaVlleu	107		
Db	15	ATCATAGAAAGTGCTG	1		
RESULT 15		CNS04L7F	1107 bp	DNA linear GSS 01-SEP-2000	
LOCUS		CNS04L7F			
DEFINITION		Tetraodon nigroviridis genome survey sequence T7 end of clone			
		118A07 of library G from Tetraodon nigroviridis, genomic survey			
		sequence.			
ACCESSION		AL295764	GI:8034344		
VERSION		GSS; genome survey sequence.			
KEYWORDS		Tetraodon nigroviridis			
SOURCE		Tetraodon nigroviridis			
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
		Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
		Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;			
		Tetraodontidae; Tetraodontinae; Tetraodon.			
REFERENCE		1 Roest Crolius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C.,			
AUTHORS		Bernot,A., Pizams,C., Winker,P., Brottier,P., Quetier,F.,			
		Saurin,W. and Weissenbach,J.			
		Estimate of human gene number provided by genome-wide analysis			
JOURNAL		using Tetraodon nigroviridis DNA sequence			
MEDLINE		Nat. Genet. 25 (2), 235-238 (2000)			
PUBMED		20296633			
REFERENCE		10835645			
AUTHORS		2 Roest Crolius,H., Jalllon,O., Dasilva,C., Ozouf-Costaz,C.,			
		Pizams,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,			
		Saurin,W., Bernot,A. and Weissenbach,J.			
TITLE		Characterization and repeat analysis of the compact genome of the			
JOURNAL		freshwater pufferfish Tetraodon nigroviridis			
MEDLINE		Genome Res. 10 (7), 939-949 (2000)			
PUBMED		20359837			
REFERENCE		10899143			
AUTHORS		3 (bases 1 to 1107)			
TITLE		Genoscope.			
JOURNAL		Direct Submission			
		Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :			
		BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)			
		- Web : www.genoscope.cns.fr			
COMMENT		This sequence is a single read and was generated as part of a large			
		scale clone-end sequencing project of the Tetraodon nigroviridis			
		genome. For more information, please take a look at			
		http://www.genoscope.cns.fr/Tetraodon.			
FEATURES		Location/Qualifiers			
source		1..1107			
		/organism="Tetraodon nigroviridis"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:99883"			
		/clone="118A07"			
		/clone_1lb="G"			
		/note="Genoscope sequence ID : COBGL18A04LP1-end : T7"			
ORIGIN					
Alignment Scores:		1.72e-15	Length:	1107	
Fred. No.:		Score:	249.00	Matches:	41
		Percent Similarity:	81.43%	Conservative:	16
		Best Local Similarity:	58.57%	Mismatches:	13
Query Match:		13.30%	Indels:	0	
DB:		9	Gaps:	0	
US-09-995-225B-16	(1-353)	x	CNS04L7F	(1-1107)	
OY	224	G UysThrThralaleleuphThrlleHseSerleephleaIathreITPalaPro	243		
Db	71	GGGAAAGGACCGGCAATGCTGCGCCATACCTCCGCTCTGTGCTGTGGACACC	130		

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 18:37:44 ; Search time 57.6327 Seconds
(without alignments)
3524.170 Million cell updates/sec

Title: US-09-995-225b-16

Perfect score: 1872

Sequence: 1 MEHTAHLAANSLSWSPG.....CIKMLVYQYDKXKPIKXSP 353

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 02.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1872	100.0	353	2 Q6DMJ6	Q6DJ66 homo sapien
2	1638	87.5	388	2 Q8TDU8	Q8TDU8 homo sapien
3	1317	70.4	265	2 Q9BYT4	Q9BYT4 homo sapien
4	1277.5	68.2	383	2 Q7T2L1	Q7T2L1 fugu rubrip
5	994	53.1	186	2 Q86SP2	Q86SP2 homo sapien
6	771	41.2	371	2 Q7T2L2	Q7T2L2 fugu rubrip
7	737	39.4	315	2 Q86SL3	Q86SL3 homo sapien
8	737	39.4	462	2 Q7Z601	Q7Z601 homo sapien
9	737	39.4	1464	2 Q8NGB0	Q8NGB0 homo sapien
10	713	38.1	365	2 Q7TGN9	Q7TGN9 mus musculi
11	262	14.0	54	2 Q80UC8	Q80UC8 mus musculi
12	260.5	13.9	394	2 Q17478	Q17478 csernorabdi
13	250.5	13.4	398	1 TRR_BOVIN	TRR_BOVIN bos taurus
14	249.5	13.3	398	1 TRR_SHEEP	TRR_SHEEP ovis aries
15	249.5	13.3	440	2 Q7PYB7	Q7PYB7 anopheles g
16	248	13.2	549	2 Q9VZM5	Q9VZM5 drosophila
17	242	12.9	412	1 TRR_RAT	TRR_RAT rattus norv
18	241	12.9	393	1 TRR_MOUSE	TRR_MOUSE mus musculi
19	239	12.8	462	2 Q8TRF3	Q8TRF3 drosophila
20	238.5	12.7	398	1 TRR_HUMAN	TRR_HUMAN homo sapien
21	238.5	12.7	398	2 AAR84356	AAR84356 homo sapi
22	236	12.6	488	2 Q717R7	Q717R7 drosophila
23	236	12.6	488	2 AAO11747	AAO11747 drosophila
24	235.5	12.6	395	1 TRR_CHICK	TRR_CHICK gallus gall
25	235	12.6	404	2 Q8JFZ7	Q8JFZ7 xenopus lae
26	234	12.5	478	2 Q717T5	Q717T5 drosophila
27	234	12.5	478	2 AAO11746	AAO11746 drosophila
28	233.5	12.5	354	2 Q6YST0	Q6YST0 sus scrofa
29	233.5	12.5	354	2 BAD12131	BAD12131 sus scrofa
30	233.5	12.5	354	2 BAD08660	BAD08660 sus scrofa
31	233	12.4	397	2 Q9DDR1	Q9DDR1 xenopus lae

32	233	12.4	540	2 Q9W025	Q9W025 drosophila
33	232	12.4	435	2 Q8SMR3	Q8SMR3 drosophila
34	231.5	12.4	355	1 CKR1_HUMAN	P32246 homo sapien
35	231.5	12.4	355	2 Q9WYJ8	Q9WYJ8 callitrich
36	231.5	12.4	355	2 AAB4991	AAB4991 homo sapi
37	231	12.3	397	2 Q8JFZ5	Q8JFZ5 xenopus lae
38	230.5	12.3	344	2 Q7PHG9	Q7PHG9 anopheles g
39	230.5	12.3	427	2 Q7RTK2	Q7RTK2 anopheles g
40	228	12.2	412	2 Q7LC54	Q7LC54 homo sapien
41	228	12.2	415	2 Q9CZQ4	Q9CZQ4 homo sapien
42	228	12.2	415	2 Q96AMS	Q96AMS homo sapien
43	228	12.2	415	2 AAB67776	AAB67776 homo sapi
44	226	12.1	352	2 Q88820	Q88820 rattus norv
45	226	12.1	352	2 Q9R297	Q9R297 rattus norv

ALIGNMENTS

RESULT 1					
Q6DMJ6	PRELIMINARY	PRT	353 AA.		
AC Q6DMJ6					
DT 01-OCT-2004 (TREMBLrel. 28, Created)					
DT 01-OCT-2004 (TREMBLrel. 28, Last sequence update)					
DT 01-OCT-2004 (TREMBLrel. 28, Last annotation update)					
DE G protein-coupled receptor 139.					
GN Name=GPR139,					
OS Homo sapiens (Human).					
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
NCBI_Taxid=9606;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC TISSUE=Brain;					
RA Bommer T.I., Nagle J.W., Kaufman D.;					
RL Submitted (MAY-2004) to the EMBL/Genbank/DBJ databases.					
DR EMBL; AY635179; AAT65818.1; -.					
KW RECEPTOR.					
SQ SEQUENCE 353 AA; 40679 MW; CF1088FC85C84F85 CRC64;					
Query Match					
Best local similarity 100.0%; Score 1872; DB 2; Length 353;					
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 MEHTAHLAANSLSWSPGACGLGFPVYVYLLLCGLPANILTVIILSQLVAROK 60				
DB	1 MEHTAHLAANSLSWSPGACGLGFPVYVYLLLCGLPANILTVIILSQLVAROK 60				
QY	61 SSYNYTLAALADIVLPFIVVDPLLEDFILNMQPOVPDKIIEVLEFSSIHSTWTV 120				
DB	61 SSYNYTLAALADIVLPFIVVDPLLEDFILNMQPOVPDKIIEVLEFSSIHSTWTV 120				
QY	121 PLTIIRYVAGPLKHTVSPARTKRVISVYITCFLSIPYWMNPWTEDYSTVA 180				
DB	121 PLTIIRYVAGPLKHTVSPARTKRVISVYITCFLSIPYWMNPWTEDYSTVA 180				
QY	181 HVLIMHCTVYLVPCSEIFILNSIIVYKLRKSNRFLNGYSTGKTALIFTTSIFATL 240				
DB	181 HVLIMHCTVYLVPCSEIFILNSIIVYKLRKSNRFLNGYSTGKTALIFTTSIFATL 240				
QY	241 WAPRIIMLYHLYGAPIONRMVLVHMSDIAANMLALNTAINFPLCFISKRRPTMAATL 300				
DB	241 WAPRIIMLYHLYGAPIONRMVLVHMSDIAANMLALNTAINFPLCFISKRRPTMAATL 300				
QY	301 KAPFKCKQKQVQFYTNHNSITSSPISPANSHCIKMLVYQYDKXKPIKXSP 353				
DB	301 KAPFKCKQKQVQFYTNHNSITSSPISPANSHCIKMLVYQYDKXKPIKXSP 353				
RESULT 2					
Q8TDU8	PRELIMINARY	PRT	388 AA.		
ID Q8TDU8					

AC Q8TDU8;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative G-protein coupled receptor.
 GN Name=GPCR;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Takeda S., Kadowaki S., Haga T., Takeasu H., Mitaku S.,
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBD databases.
 DR EMBL; AB083594; BAB89307.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GFCRRHODPSN.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
 KW Receptor.
 SQ SEQUENCE 388 AA; 44893 MW; D8DF1CF0CC95D9D8 CRC64;

Query Match 87.5%; Score 1638; DB 2; Length 388;
 Best Local Similarity 100.0%; Pred. No. 1.9e-105;
 Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 ANILTVIIISQLVARQKSSNYLLAALAAADILVFFIVFDLLEDFILNMQMPQVDPK 102
 DB 78 ANILTVIIISQLVARQKSSNYLLAALAAADILVFFIVFDLLEDFILNMQMPQVDPK 137
 QY 103 IIEVLEFSSHTSIWITVPLTIDRYAVCHPLKHTVSPARTRKVIYSVYITCFLTSLIP 162
 DB 138 IIEVLEFSSHTSIWITVPLTIDRYAVCHPLKHTVSPARTRKVIYSVYITCFLTSLIP 197
 QY 163 YWMWNTEDYISTSVHVLWVHCFVYVLPVCSIFPLINSIIYKLRKSNFPLRGYS 222
 DB 198 YWMWNTEDYISTSVHVLWVHCFVYVLPVCSIFPLINSIIYKLRKSNFPLRGYS 257
 QY 223 TGTATLIFTTISIPATLAPRIIMILYHLYGAPLQNRWLVIHMSDIANMLALTANIN 282
 DB 258 TGTATLIFTTISIPATLAPRIIMILYHLYGAPLQNRWLVIHMSDIANMLALTANIN 317
 QY 263 FLYCFISKRFRMAATLKAFKCKQKQPVQFTNNHFSITSSPWISPCISKMLVYQY 342
 DB 318 FLYCFISKRFRMAATLKAFKCKQKQPVQFTNNHFSITSSPWISPCISKMLVYQY 377
 QY 343 DKGKPKIKVSP 353
 DB 378 DKGKPKIKVSP 388

RESULT 3

Q9BYT4 PRELIMINARY; PRT; 265 AA.
 AC Q9BYT4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative transmembrane receptor (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21853304; PubMed=11863363;
 RA Octolenghi C., Barbieri M., McElwee K., Fellous M.,
 RT "Novel paralogy relations among human chromosomes support a link
 RT between the phylogeny of doublesex-related genes and the evolution of

RT sex determination.";
 RL Genomics 79:333-343(2002).
 DR EMBL; AJ303165; CAC33085.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GFCRRHODPSN.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
 KW Receptor; Transmembrane.
 FT NON TER 1
 FT NON TER 1
 SQ SEQUENCE 265 AA; 30999 MW; 8F314368A492395B CRC64;

Query Match 70.4%; Score 1317; DB 2; Length 265;
 Best Local Similarity 100.0%; Pred. No. 2.1e-83;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 ANILTVIIISQLVARQKSSNYLLAALAAADILVFFIVFDLLEDFILNMQMPQVDPK 102
 DB 14 ANILTVIIISQLVARQKSSNYLLAALAAADILVFFIVFDLLEDFILNMQMPQVDPK 73
 QY 103 IIEVLEFSSHTSIWITVPLTIDRYAVCHPLKHTVSPARTRKVIYSVYITCFLTSLIP 162
 DB 74 IIEVLEFSSHTSIWITVPLTIDRYAVCHPLKHTVSPARTRKVIYSVYITCFLTSLIP 133
 QY 163 YWMWNTEDYISTSVHVLWVHCFVYVLPVCSIFPLINSIIYKLRKSNFPLRGYS 222
 DB 134 YWMWNTEDYISTSVHVLWVHCFVYVLPVCSIFPLINSIIYKLRKSNFPLRGYS 193
 QY 223 TGTATLIFTTISIPATLAPRIIMILYHLYGAPLQNRWLVIHMSDIANMLALTANIN 282
 DB 194 TGTATLIFTTISIPATLAPRIIMILYHLYGAPLQNRWLVIHMSDIANMLALTANIN 253
 QY 283 FLYCFISKRFR 294
 DB 254 FLYCFISKRFR 265

RESULT 4

Q7T2LI PRELIMINARY; PRT; 383 AA.
 ID Q7T2LI
 AC Q7T2LI;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE G protein-coupled receptor 142b (Fragment).
 GN Name=GPCR142b;
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Takifugu.
 OX NCBI_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22985413; PubMed=14623098;
 RA Fredriksson R., Hoglund P.O., Gloriam D.E., Lagerstrom M.C.,
 RA Schioth H.B.;
 RT "Seven evolutionarily conserved human rhodopsin G protein-coupled
 RT receptors lacking close relatives."
 RL FEBS Lett. 554:381-388(2003).
 DR EMBL; AY288414; AAP72123.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_FL_2; 1.
 KW Receptor.

FT NON TER 1 1
 SQ SEQUENCE 383 AA; 42610 MW; 761B0AA3B97ECDD1 CRC64;

Query Match 68.2%; Score 1277.5; DB 2; Length 383;
 Best Local Similarity 66.3%; Pred. No. 1.6e-80;

Matches 250; Conservative 35; Mismatches 67; Indels 25; Gaps 5;

QY 1 MHTHTAHLANSLSLWSPG-----SACGLGFVVVVVYLLCLGLP-ANLITVYIIS 52
 DB 6 MHSHTFSALEPSSWSSTGQHPSEATQCCPLPLPIYSSALCLGLPVALNLTIVYIS 65
 QY 53 QLVARRKSSSYNYLLAALADILVLFPIFVDFLEDFILNMQPOVPDKIEVLEFSST 112
 DB 66 QLVMRKSSSYNYLLAALADILVLLIFVDFLEDFILATLPPLSNAVQVLEFSST 125
 QY 113 HTSIWITVPLTIDRYAVCHPLKHTVSPARTKXIVSVYITCFILTSIPYWMPIWTE 172
 DB 126 HTSIWITVPLTIDRYAVCHPLKHTVSPARTKXIVSVYITCFILTSIPYWMPIWTE 185
 QY 173 -----DYISTVHNLWICHTVYLVCSIFPILNSTITVYKL-RRKSNPL 218
 DB 186 LPQTGGGGAGGAGGNRRVTAQOILVMAHCITVYFLCTVFFSLNAVYVHKLRRRSCFRL 245
 QY 219 RGYSTKTALLETTSIFATLWAPRIIMLYHLYGAPIQNRW---LVHIMSDIANMLAL 275
 DB 246 RGYSTKTALLETTSIFATLWAPRIIMLYHLYGAPIQNRW---LVHIMSDIANMLAL 305
 QY 276 LNTAINFLYCFISKFRMTAAATLKAFFKOKOPVOPTNNHNSITSSPWISPA NSHCT 335
 DB 306 LNTGVNPLFYCFISKFRMGAAVNLALVCRKQOPFVASHNFSITSSPWISPA NSHCT 365
 QY 336 KMLVVOYDKNGKPKIKVS 352
 DB 366 KMLVVOYDKNGKPKIKVS 382

RESULT 5

Q66SP2 PRELIMINARY; PRT; 186 AA.
 AC Q66SP2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE G protein-coupled receptor FGR3 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NC NCB1_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22584407; PubMed=12679517;
 RA Vasellakis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
 Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
 RA Bergmann J.E., Gallanaris G.A.;
 RT "The G protein-coupled receptor repertoire of human and mouse."
 RT Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003).
 DR EMBL; AY255545; AA085057.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsin.
 DR PRINTS; PR00237; GPCR_Rhodopsin.
 DR PROSITE; PS50262; G_PROTEIN_RECPEP_F1_2; 1.
 DR RECEPTOR.
 FT NON TER 1 1
 SQ SEQUENCE 186 AA; 21687 MW; E402AB947EC080AF CRC64;

Query Match 53.1%; Score 994; DB 2; Length 186;
 Best Local Similarity 100.0%; Pred. No. 3.3e-61;
 Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 NIWTEDYISTVHNLWICHTVYLVCSIFPILNSTITVYKLRRKSNPLRGYSTGKT 227
 DB 1 NIWTEDYISTVHNLWICHTVYLVCSIFPILNSTITVYKLRRKSNPLRGYSTGKT 60
 QY 228 AILFTTTSIFATLWAPRIIMLYHLYGAPIQNRWLVHIMSDIANMLALNTAINPLVYCF 287
 DB 61 AILFTTTSIFATLWAPRIIMLYHLYGAPIQNRWLVHIMSDIANMLALNTAINPLVYCF 120
 QY 288 ISKRFRTMAATLKAFFKOKOPVOPTNNHNSITSSPWISPA NSHCTKMLVVOYDKNGK 347
 DB 121 ISKRFRTMAATLKAFFKOKOPVOPTNNHNSITSSPWISPA NSHCTKMLVVOYDKNGK 180
 QY 348 PIKVPSP 353
 DB 181 PIKVPSP 186

RESULT 6

Q72L2 PRELIMINARY; PRT; 371 AA.
 AC Q72L2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE G protein-coupled receptor 142a (Fragment).
 GN Name=GR142a;
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Acanthopterygii; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Takifugu.
 NC NCB1_Taxid=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22985413; PubMed=14623098;
 RA Fredriksson R., Hoglund P.J., Gloriam D.E., Lagerstrom M.C.,
 RA Schloth H.B.;
 RT "Seven evolutionarily conserved human rhodopsin G protein-coupled
 RT receptors lacking close relatives."
 RL FEBS Lett. 554:381-388(2003).
 DR EMBL; AY288413; AAP72122.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsin.
 DR PROSITE; PS50237; G_PROTEIN_RECPEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECPEP_F1_2; 1.
 DR RECEPTOR.
 FT NON TER 1 1
 SQ SEQUENCE 371 AA; 42000 MW; C1A95B6619681AB3 CRC64;

Query Match 41.2%; Score 771; DB 2; Length 371;
 Best Local Similarity 47.0%; Pred. No. 1.6e-45;
 Matches 133; Conservative 65; Mismatches 75; Indels 10; Gaps 2;

Query Match	39.4%	Score 737;	DB 2;	Length 1464;
Best Local Similarity	46.3%	Pred. No. 1.2e-42;		
Matches 133;	Conservative 57;	Mismatches 95;	Indels 2;	Gaps 2
DR	GO: 0001584; F: rhodopsin-like receptor activity; IEA.			
DR	GO: 0007166; P: G-protein coupled receptor protein signaln. . .; IEA.			
DR	InterPro; IPR000276; GPCR_Rhodopsn.			
DR	Pfam; PF00001; 7tm_1; 1.			
DR	PRINTS; PR00237; GPCR_RHODOPSIN.			
DR	PROSITE; PS00237; G PROTEIN RECP_F1_1; UNKNOWN_1.			
DR	PROSITE; PS00262; G PROTEIN RECP_F1_2; 1.			
KW	Receptor; Transmembrane			
SO	SEQUENCE 1464 AA; 156507 MW; FFGCG3F835ACF7E CRC64;			
Query Match	39.4%	Score 737;	DB 2;	Length 1464;
Best Local Similarity	46.3%	Pred. No. 1.2e-42;		
Matches 133;	Conservative 57;	Mismatches 95;	Indels 2;	Gaps 2
DR	GO: 0001584; F: rhodopsin-like receptor activity; IEA.			
DR	GO: 0007166; P: G-protein coupled receptor protein signaln. . .; IEA.			
DR	InterPro; IPR000276; GPCR_Rhodopsn.			
DR	Pfam; PF00001; 7tm_1; 1.			
DR	PRINTS; PR00237; GPCR_RHODOPSIN.			
DR	PROSITE; PS00237; G PROTEIN RECP_F1_1; UNKNOWN_1.			
DR	PROSITE; PS00262; G PROTEIN RECP_F1_2; 1.			
KW	Receptor; Transmembrane			
SO	SEQUENCE 1464 AA; 156507 MW; FFGCG3F835ACF7E CRC64;			

[illegible]

RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RL investigating biology".
 RL Science 282:2012-2018 (1998).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Brisc1 N2;
 RA Swindburne J.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z66519; CA91374.3; -.
 DR PIR: T18704; T18704.
 DR WormPep; B0334.6; CE30473.
 DR GO; GO:0016021; C: integral to membrane, IEA.
 DR GO; GO:0001584; F: rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P: G-protein coupled receptor protein signaling. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOPOPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 394 AA; 45120 MW; E0EC618D653B8627 CRC64;
 Query Match 13.4%; Score 260.5; DB 2; Length 394;
 Best Local Similarity 23.6%; Pred. No. 3.4e-10;
 Matches 90; Conservative 62; Mismatches 151; Indels 79; Gaps 12;
 QY 5 HAHLANSSLSWMSPGSACGLGVPPVYVYSLLCGLPANILTVIISQVLRQKSSYN 64
 DB 31 HEHDEQSIIVWMS-----NVAVLVYI-----ALIGLACNLMMVLTSNKTRRIPSWN 80
 QY 65 YLLAIAADILVLFVFDVFLLEDFILNMOMPOVD-----KIEVLEF 109
 DB 81 LLIALAVCSLFLIFAT-----LDVTELSIPSLAFSTSFNFYSRIYLRITLAS 130
 QY 110 SSIHNSIWTFPLTDRIYAVCHPLKHTVSPARTKRYVSVVYTCFLTSIPYWWP-- 167
 DB 131 TTYKSSVLLVAVFNIERLYCVVCPINSHRWCTSRNSKAIATAYLSFLCSIQ--WPLA 187
 QY 168 --NWTEDYISTSVHHVLI-----WICHTVYLVCSIFILNLSIIVKL 210
 DB 188 YPTICFESNSQYVYVILMSTRALQIYRTMDVSLPAFNVLPITGLVYNSRIIFTL 247
 QY 211 RKSNFRLRGYS-----TGKTAILEFTISIFATLMAPRI--IMILYHL 252
 DB 248 RRVVDESRKYETKSLDGLIQDAHNRTRBANMLFAVYVWMLFCGQAPARILVDM 307
 QY 253 YGAPLQNMVLIHIMEDIANMLALTAINFLYCFISGRFTMAATLKAF--KQKQPV 311
 DB 308 YG-QYHFAILIVC--LSQQLVFLNASINFLCYCVVSKRYRTLMKQTLKKFLHKLGVGH 364
 QY 312 QYTNHNSITSSPWISPSNSH 333
 DB 365 PFQINLKQTKSSSAHTSLEDH 386
 RESULT 13
 TRFR BOVIN STANDARD; PRT; 398 AA.
 ID AC 046639;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin
 receptor).
 GN Name=TRHR;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=Holstein;
 RX MEDLINE=9815137; PubMed=9492373;
 RA Taketa M., Shimada Y., Ikeda A., Sekikawa K.;
 RT "Molecular cloning of bovine thyrotropin-releasing hormone receptor
 gene".
 RL J. Vet. Med. Sci. 60:123-127 (1998).
 CC -!- FUNCTION: Receptor for thyrotropin-releasing hormone. This
 CC receptor is mediated by G proteins which activate a
 CC phosphatidylinositol-calcium second messenger system.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; D83964; BA24069.1; -.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR009144; ThyrotropinRHR.
 DR InterPro; IPR002120; Trelhnm_receptor.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOPOPSN.
 DR PROSITE; PR01846; TRHRFAMILY.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1;
 DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 DR G-protein coupled receptor; Glycoprotein; Transmembrane.
 FT DOMAIN 1 28
 FT TRANSMEM 29 51
 FT DOMAIN 52 61
 FT TRANSMEM 62 83
 FT DOMAIN 84 99
 FT TRANSMEM 100 121
 FT DOMAIN 122 144
 FT TRANSMEM 145 168
 FT DOMAIN 169 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 266
 FT TRANSMEM 267 288
 FT DOMAIN 289 296
 FT TRANSMEM 297 319
 FT DOMAIN 320 398
 FT DISULFID 98 179
 FT CARBOHYD 3 3
 FT CARBOHYD 10 10
 SQ SEQUENCE 398 AA; 45165 MW; D9AF4B211A5701B8 CRC64;
 Query Match 13.4%; Score 250.5; DB 1; Length 398;
 Best Local Similarity 23.2%; Pred. No. 1.7e-09;
 Matches 83; Conservative 69; Mismatches 118; Indels 87; Gaps 13;
 QY 28 VPVVYVSYLLC-LGLPANILTVIISQVLRQKSSYN-YLLAIAADILVLFVFDV 85
 DB 26 VVITLAVLIIIGLGIYGNIMVVLVWNR--TKMKRTPNCTIVSLAADLMVL----- 75
 QY 86 LLEDFILNMOMPOVPDKI-----IEVLEFSSIHSTWTFPLTDRIYAVCH 132
 DB 76 -----VAAGLPNITDSIYGSWYGVGCLITYLGLGNASSCSITAFTRRYIAICH 129
 QY 133 PLKHYVSYARFRKRYVSVYITCFILTSIPYWW-----PNWT-EDYSTSVHHVLI 186
 DB 130 PIRAGPLCTFSRKKIIFVWA--FTSIYCMFLDLNLTSTYDAIVVSGYKISR 186
 QY 187 HCFVYLVCSIFPLINSIIVKL-----RKSN----- 215
 DB 187 YSPPIYLMRGVYVPMILATVLYGFARILFLANIPSDPKENSMKKDSTHQNKN 246
 QY 216 -----FLRGYSTKTTALIFTTISIFATLMAP--RIIMLYHLGAPLQNMVLIH 267
 DB 247 SKTSNRYFNSVSSRKQVTKMLAVVILPALIMMPRTLVVNSFLSSPFQENWFL 303

QY 268 DIANMALNTAINTFLYCFISKRTMAATLKAFKCKOKOPVOFTYTHNSITS 324
 DB 304 -FCRICIYLSAINPVIYNLMGOKFR---AARFKLCNCKOKRVEKPNYSVALNS 355

RESULT 14
 TRPR_SHEEP STANDARD, PRT, 398 AA.

AC Q28596, 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin receptor).
 GN Name=TRHR;
 OS Ovis aries (Sheep); Chordata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Caprinae; Ovis;
 NCBI_TaxId=9940;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Plutitary;
 RX MEDLINE=97200775; PubMed=9048604;
 RA Bockman J., Boeckers T.M., Winter C., Wiltkowsk W., Winterhoff H., Deutel T., Kreutz M.R.;
 RT "Thyrotropin expression in hypophyseal pars tuberalis-specific cells is 3,5',5'-triodothyronine, thyrotropin-releasing hormone, and pit-1 independent.";
 RT Endocrinology 138:1019-1028(1997).
 CC -1- FUNCTION: Receptor for thyrotropin-releasing hormone. This receptor is mediated by G proteins which activate a phosphatidylinositol-calcium second messenger system.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X95285; CAA64606.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR009144; ThyrotropinRHR.
 DR InterPro: IPR002120; Trh1hm_receptor.
 DR Pfam: PF00001; 7tm_1.1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PRO1846; TRHRFAMILY.
 DR PROSITE: PS00237; G PROTEIN RECP F1.1; 1.
 DR PROSITE: PS00262; G PROTEIN RECP F1.2; 1.
 KM G-protein coupled receptor; Glycoprotein; Transmembrane. Extracellular (Potential).
 FT DOMAIN 1 28
 FT TRANSMEM 29 51
 FT DOMAIN 52 61
 FT TRANSMEM 62 83
 FT DOMAIN 84 99
 FT TRANSMEM 100 121
 FT DOMAIN 122 144
 FT TRANSMEM 145 168
 FT DOMAIN 169 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 266
 FT TRANSMEM 267 288
 FT TRANSMEM 289 296
 FT TRANSMEM 297 319
 FT TRANSMEM 320 398
 FT DISULFID 98 179
 FT CARBOHYD 3 3
 FT CARBOHYD 10 10
 N-linked (GlcNAc. .) (Potential).
 N-linked (GlcNAc. .) (Potential).

SQ SEQUENCE 398 AA; 45088 MW; 375A311D3DD2A61A CRC64;
 Query Match 13.3%; Score 249.5; DB 1; Length 398;
 Best Local Similarity 23.2%; Pred. No. 1.9e-09;
 Matches 83; Conservative 68; Mismatches 119; Indels 87; Gaps 13;

QY 28 VVVVYSLDLC-IGPANITVITISQVARRKSSYN-YLDAALADILVFFIVDF 85
 DB 26 VVTILLVITICGLGIVGINMVLVVR--TKMRPTNCTIVSLAVADMVL----- 75
 QY 86 LLEDFLLNMQMPQVDPKI-----IEVLEPSSIHITSIMITVPLTIDRYAVCH 132
 DB 76 -----VAAGLBNITDSIGSWYGVGCLCTIYQYIGINSSCSITAFTERIYAICH 129
 QY 133 PLKHTVSPARTKRVYSVYITCTPLTSPYTW-----PNIWT-EDYISTSVHVLIM 186
 DB 130 PIKAQFLCTPSRAKIIIFVWA--FTSYICMLMFFLLDLNISTYKDAIVSCGYKISN 186
 QY 187 HCFVYLVPCSIFFLINSIIVYKL-----RKSN----- 215
 DB 187 YSPYIYNDPGVYFVVPMLATVYGFARILFLSPIDPDKENSTWKNDSHTONKLN 246
 QY 216 -----FRLRGYSTGKTTAIFITTSIPATLMAR-RIIMIDYHVGADIQRMVLVHNS 267
 DB 247 SKTSNRYFNSTVSSRKQVTKMLAVVILFALIMPFRTLVVNSFLSPQENWFL--- 303
 QY 268 DIANMALNTAINTFLYCFISKRTMAATLKAFKCKOKOPVOFTYTHNSITS 324
 DB 304 -FCRICIYLSAINPVIYNLMGOKFR---AARFKLCNCKOKRVEKPNYSVALNS 355

RESULT 15
 QY/PYB7 PRELIMINARY, PRT, 440 AA.

AC Q7PYB7;
 DT 01-MAR-2004 (TREMBLrel. 26, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE AGCP12601.
 GN Name=agCG53608; ORFName=ENSANG0000018385;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
 NCBI_TaxId=180454;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
 CC EMBL: AAB01008987; EAA01756.1; -
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1.1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00262; G PROTEIN RECP F1.2; 1.
 SQ SEQUENCE 440 AA; 50079 MW; 953A62F3E6A9C95 CRC64;

Query Match 13.3%; Score 249.5; DB 2; Length 440;
 Best Local Similarity 23.3%; Pred. No. 2.1e-09;
 Matches 69; Conservative 67; Mismatches 115; Indels 45; Gaps 8;

QY 31 VVYSLDLCIGPANITVITISQVARRKSSYNLT-ALAAADILVFFIVDFLEED 89
 DB 35 VVMTVALIGLIGNIFSWILSR---PQRRSINVLGLACDVLITSLTSLIGLCAI 91
 QY 90 F-----ILNMQMPQVDPKI---IEVLEPSSIHITSIMITVPLTIDRYAVCHPLKHTVSY 142
 DB 92 YHTGVLYVYHNIQIPKISLVVYPLAMTAQNTASVLTITLTERVAVACHPFRARALCTY 151

Wed Dec 15 11:04:28 2004

us-09-995-225b-16.rup

Page 8

```

OY 143 ARTRKAVSYITTCELTSPFYWMENI-----WTEDYSTSVHH 161
Db 152 GRALYYVGVILVSIILNLEPFMEVYLLISTHPDTGLTYCYKASDMKNTETTYIKVYIH- 210
OY 182 VLIMHCFEYVLPSCSFPIINSLIYYKLRKRSNRLGYSGTG---TAILFTINSIP 237
Db 211 ---WLMVFIFYPFPLISEFNMIMYRÖVRANKROBLSRSEKREIGLMTLCCVIVF 267
OY 238 ATLMAPIIMILYHLYGADIQNRMLVHIMSDIANMLALNPAINEPLCYFISIKSR 293
Db 268 LLONLPAAMINIVEAFYSVI---IEYVVKSNLVTINSVAFETIYFGCEFFK 318

```

Search completed: December 14, 2004, 20:25:08
Job time : 59.6327 secs

Db 181 ATGCCGAGCTGTACTTTGTCAATGGCAGTGGACGCGCTGTGCTCAGCGCCCTGGCC 240
QY 241 CCTGTGCACTGTCTGGCCCCCGAGCTCCGGTGGGCGCTGTGAAGTGGGCGGCGAA 300
Db 241 CCTGTGCACTGTCTGGCCCCCGAGCTCCGGTGGGCGCTGTGAAGTGGGCGGCGAA 300
QY 301 GTTCCAGTGGCACTGAGAGATCCCTTCAATGTGTCTCACTGTGGCCATGTAATCCAC 360
Db 301 GTTCCAGTGGCACTGAGAGATCCCTTCAATGTGTCTCACTGTGGCCATGTAATCCAC 360
QY 361 GCCCTGTGAGCTTCGACCACTACATGAGCGCTGCTCCGCGACCTACATGCGCAC 420
Db 361 GCCCTGTGAGCTTCGACCACTACATGAGCGCTGCTCCGCGACCTACATGCGCAC 420
QY 421 GTGTACACACGCGGACGCTGTGGGCTTGTGTGGGCTGTGGCGCTGTGACCACTTC 480
Db 421 GTGTACACACGCGGACGCTGTGGGCTTGTGTGGGCTGTGGCGCTGTGACCACTTC 480
QY 481 TCTCTGCTGTCTTTCAATCTGACACCATGTGTCCACCCGCGCGCTAGAGTGGCGAAG 540
Db 481 TCTCTGCTGTCTTTCAATCTGACACCATGTGTCCACCCGCGCGCTAGAGTGGCGAAG 540
QY 541 ATSCAGAACGCAAGACTCCGACGCGACGCTGTGTATCGGCTACGCTAGTGGCCAGA 600
Db 541 ATSCAGAACGCAAGACTCCGACGCGACGCTGTGTATCGGCTACGCTAGTGGCCAGA 600
QY 601 CTGGCCACCTCTTACGCGCTGTGTACTCTCCGCTCCGACGAGGAGACACGCGCTTG 660
Db 601 CTGGCCACCTCTTACGCGCTGTGTACTCTCCGCTCCGACGAGGAGACACGCGCTTG 660
QY 661 GACCGGAGACACGCGGCGCGCTGGAGCCCTGGGACACAGAGCTGTGTGGCGACCGTGTG 720
Db 661 GACCGGAGACACGCGGCGCGCTGGAGCCCTGGGACACAGAGCTGTGTGGCGACCGTGTG 720
QY 721 ACCGAGTTTGGGCTGTGACGCGCACTATCTGATCTGTGGGACACAGGTCATCATC 780
Db 721 ACCGAGTTTGGGCTGTGACGCGCACTATCTGATCTGTGGGACACAGGTCATCATC 780
QY 781 TCCGAGGAGGAGCCCGTGGACGCACTACTTGGGCTGTGCACTTTGTGAAGGATTTTC 840
Db 781 TCCGAGGAGGAGCCCGTGGACGCACTACTTGGGCTGTGCACTTTGTGAAGGATTTTC 840
QY 841 TCCAACTCTGTGGCTTCTCAGAGCTTGTGACACACTTCTCTACGCTACATGAAC 900
Db 841 TCCAACTCTGTGGCTTCTCAGAGCTTGTGACACACTTCTCTACGCTACATGAAC 900
QY 901 CAGAGCTTCCCGACAGACTCCAAAGCTGTGATGAAAAAGCTGCGCGGAGACGGAC 960
Db 901 CAGAGCTTCCCGACAGACTCCAAAGCTGTGATGAAAAAGCTGCGCGGAGACGGAC 960
QY 961 TGTCTCCCGGACCAATGGGGGTGACAGAGTGTGCGGTAG 1002
Db 961 TGTCTCCCGGACCAATGGGGGTGACAGAGTGTGCGGTAG 1002

RESULT 2
AX665926 1002 bp DNA linear PAT 26-MAR-2003
LOCUS AX665926 Sequence 3 from Patent WO242458.
DEFINITION AX665926
ACCESSION AX665926
VERSION AX665926.1 GI:29290799
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Tian,H., Zhao,J., Chen,J.L., Cutler,G., An,S., Dai,K. and Gupta,J.S.
AUTHORS G-protein coupled receptors
TITLE Patent: WO 0242458-A 3 30-MAY-2002;
JOURNAL Tularik Inc. (US)
FEATURES Location/Qualifiers

source 1..1002
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 1..1002
/note="unassigned protein product; human G-protein coupled receptor (GPCR) TGR339"
/codon_start=1
/protein_id="CAD82888.1"
/db_xref="GI:29290800"
/translation="MMSCSMFPNGTGLVEELPACODLQGLSLSLGLVGVYGLCY NALVIVANLHASKAMTMDVYFVMAVAGVLSALAPVHLGLPSSRWALMSVGEVH VALDIPFVSLVMAVSTALSLDHYERLAPRYMASVYVTRHVCGRWALSTF SLLFVYCSHVSSTRALRCAKQWNAEADATLVPFYVPAVATLALVLSRVREDT PLDDTGTLEBSARLLVATCTGPGMLPHYLILGLTVLISGRKPDADHYLGLHF VKDPSKLIAFSFSFVTLVLRVNMQSPFSKLQRLMKDPCGDRHCSPDHMQVYLA"
ORIGIN
Query Match 100.0%; Score 1002; DB 6; Length 1002;
Best Local Similarity 100.0%; Pred. No. 3,66-163;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGAGCTGAGCTGTGTTCAACGCGACAGGCGCTGTGTGAGAGAGCTGCTGCTGAG 60
Db 1 ATGTGAGCTGAGCTGTGTTCAACGCGACAGGCGCTGTGTGAGAGAGCTGCTGCTGAG 60
QY 61 GACCTGACGCTGGGCTGTGCACTGTGTGTGCTGTGGGCTGTGTGTGGGCGTGCAGTG 120
Db 61 GACCTGACGCTGGGCTGTGCACTGTGTGTGCTGTGGGCTGTGTGTGGGCGTGCAGTG 120
QY 61 GACCTGACGCTGGGCTGTGCACTGTGTGTGCTGTGGGCTGTGTGTGGGCGTGCAGTG 120
Db 61 GACCTGACGCTGGGCTGTGCACTGTGTGTGCTGTGGGCTGTGTGTGGGCGTGCAGTG 120
QY 121 GGCCTGTGCTCAACAGCCCTGCTGTGTGCGCACTTACACAGCAAGCAGCATGAC 180
Db 121 GGCCTGTGCTCAACAGCCCTGCTGTGTGCGCACTTACACAGCAAGCAGCATGAC 180
QY 121 GGCCTGTGCTCAACAGCCCTGCTGTGTGCGCACTTACACAGCAAGCAGCATGAC 180
Db 121 GGCCTGTGCTCAACAGCCCTGCTGTGTGCGCACTTACACAGCAAGCAGCATGAC 180
QY 181 ATGCCGAGCTGTACTTTGTCAACATGGAGTGGAGGCTGTGTCTAGGCGCTTGACC 240
Db 181 ATGCCGAGCTGTACTTTGTCAACATGGAGTGGAGGCTGTGTCTAGGCGCTTGACC 240
QY 181 ATGCCGAGCTGTACTTTGTCAACATGGAGTGGAGGCTGTGTCTAGGCGCTTGACC 240
Db 181 ATGCCGAGCTGTACTTTGTCAACATGGAGTGGAGGCTGTGTCTAGGCGCTTGACC 240
QY 241 CCTGTGACCTGTGTGGGCGCGCGAGCTCCGGTGGGCGCTGTGAGTGGGCGGCGAA 300
Db 241 CCTGTGACCTGTGTGGGCGCGCGAGCTCCGGTGGGCGCTGTGAGTGGGCGGCGAA 300
QY 241 CCTGTGACCTGTGTGGGCGCGCGAGCTCCGGTGGGCGCTGTGAGTGGGCGGCGAA 300
Db 241 CCTGTGACCTGTGTGGGCGCGCGAGCTCCGGTGGGCGCTGTGAGTGGGCGGCGAA 300
QY 301 GTTCAAGTGGCACTGACAGATCCCTTCAATGTGTCTCACTGTGTGGCATGTCTCAC 360
Db 301 GTTCAAGTGGCACTGACAGATCCCTTCAATGTGTCTCACTGTGTGGCATGTCTCAC 360
QY 361 GCGCTGTGAGCTTGCACCACTACATGAGCGTGTACTGCGCGGACCTTACATGCGCAC 420
Db 361 GCGCTGTGAGCTTGCACCACTACATGAGCGTGTACTGCGCGGACCTTACATGCGCAC 420
QY 361 GCGCTGTGAGCTTGCACCACTACATGAGCGTGTACTGCGCGGACCTTACATGCGCAC 420
Db 361 GCGCTGTGAGCTTGCACCACTACATGAGCGTGTACTGCGCGGACCTTACATGCGCAC 420
QY 421 GTGTACACACGCGGACGCTGTGGGCTTGTGTGGGCTGTGGCGCTGTGACCACTTC 480
Db 421 GTGTACACACGCGGACGCTGTGGGCTTGTGTGGGCTGTGGCGCTGTGACCACTTC 480
QY 481 TCTCTGCTGTCTTTCAATCTGACACCATGTGTCCACCCGCGCGCTAGAGTGGCGAAG 540
Db 481 TCTCTGCTGTCTTTCAATCTGACACCATGTGTCCACCCGCGCGCTAGAGTGGCGAAG 540
QY 481 TCTCTGCTGTCTTTCAATCTGACACCATGTGTCCACCCGCGCGCTAGAGTGGCGAAG 540
Db 481 TCTCTGCTGTCTTTCAATCTGACACCATGTGTCCACCCGCGCGCTAGAGTGGCGAAG 540
QY 541 ATSCAGAACGCAAGACTCCGACGCGACGCTGTGTATCGGCTACGCTAGTGGCCAGA 600
Db 541 ATSCAGAACGCAAGACTCCGACGCGACGCTGTGTATCGGCTACGCTAGTGGCCAGA 600
QY 541 ATSCAGAACGCAAGACTCCGACGCGACGCTGTGTATCGGCTACGCTAGTGGCCAGA 600
Db 541 ATSCAGAACGCAAGACTCCGACGCGACGCTGTGTATCGGCTACGCTAGTGGCCAGA 600
QY 601 CTGGCCACCTCTTACGCGCTGTGTACTCTCCGCTCCGACGAGGAGACACGCGCTTG 660
Db 601 CTGGCCACCTCTTACGCGCTGTGTACTCTCCGCTCCGACGAGGAGACACGCGCTTG 660
QY 661 GACCGGAGACACGCGGCGCGCTGGAGCCCTGGGACACAGAGCTGTGTGGCGACCGTGTG 720
Db 661 GACCGGAGACACGCGGCGCGCTGGAGCCCTGGGACACAGAGCTGTGTGGCGACCGTGTG 720
QY 721 ATSCAGTTTGGGCTGTGACGCGCACTATCTGATCTGTGGGCGACAGGTCATCATC 780
Db 721 ATSCAGTTTGGGCTGTGACGCGCACTATCTGATCTGTGGGCGACAGGTCATCATC 780

QY 781 TCGCGAGGAGAGCCCGTGGACGACACTACCTGGGGCTACTGCACTTTTGAAGATTTT 840
DB 781 TCGCGAGGAGAGCCCGTGGACGACACTACCTGGGGCTACTGCACTTTTGAAGATTTT 840
QY 841 TCCAAACTCTGGCTCTTCTCCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
DB 841 TCCAAACTCTGGCTCTTCTCCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
QY 901 CAGAGCTTCCCGACGACAGCTCCCAACGCGTGATGAAAAAGCTGCCGCGGAGCCGAC 960
DB 901 CAGAGCTTCCCGACGACAGCTCCCAACGCGTGATGAAAAAGCTGCCGCGGAGCCGAC 960
QY 961 TCGTCCCGGACCACTAGGGGGGTGACGAGCTGCTGGCGGTAG 1002
DB 961 TCGTCCCGGACCACTAGGGGGGTGACGAGCTGCTGGCGGTAG 1002

RESULT 3
AX646823 1402 bp DNA linear PAT 04-MAR-2003
LOCUS AX646823
DEFINITION Sequence 1015 from Patent EP1270724.
ACCESSION AX646823
VERSION AX646823.1 GI:28799247
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.
TITLE Guanosine triphosphate-binding protein coupled receptors
JOURNAL Patent: EP 1270724-A 1015 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology
(JIP) / Center for Advanced Science and Technology Incubation, Ltd.
(JIP)

FEATURES
source location/Qualifiers
1..1402
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
201..1202
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD69537.1"
/db_xref="GI:28799248"
/translation="MNSCSWFMNGTGLVBEIPACODQLGSLSLGLVGVVGVGLCY
NALPLANLHSHKSMTPDVPYFVNMVAVGLVSLAPVHLGPPSRGMLMVSQGEVH
VLAQIPFNVSILVAMYSTALLSLDHYIBALPRTVAVSVNTRHVGSPFMWGGALITSP
SSLIPTICSHVSTRALAECAKMQNABADATLIVPTIGVVALATLYALVLSRYRREDT
PLDRDTGRLEPSSAHLVAVTCTQFGLMTPHYILIGRTVILSKRPVDATYGLLHF
VDFSKLLAFSSSFVPLLYRYNOSFPSKLRIMKRLPCGDHSCSPDHMGVOOVLAA"

ORIGIN
Query Match 100.0%; Score 1002; DB 6; Length 1402;
Best Local Similarity 100.0%; Pred. No. 3.5e-163; Indels 0; Gaps 0;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGACAGCTGTTTCAACGACACAGGCTGGTGAAGAGCTGCTGCTCCAG 60
DB 201 ATGTGAGCTGACAGCTGTTTCAACGACACAGGCTGGTGAAGAGCTGCTGCTCCAG 60
QY 61 GACCTGACAGCTGGGCTGTCTACTGTTGCTGCTGCTGGGCTGGTGGGCGTCCAGT 120
DB 261 GACCTGACAGCTGGGCTGTCTACTGTTGCTGCTGCTGGGCTGGTGGGCGTCCAGT 120
QY 121 GGGCTGTGACAAAGCCCTGTGCTGGTGGGCACTTACACAGCAAGGCGACATGACC 180
DB 321 GGGCTGTGACAAAGCCCTGTGCTGGTGGGCACTTACACAGCAAGGCGACATGACC 180
QY 181 ATGCCGAGACGTACTTGTTCACATGACAGTGGACGCGCTGGTGTCAAGCGCTGGCC 240
DB 381 ATGCCGAGACGTACTTGTTCACATGACAGTGGACGCGCTGGTGTCAAGCGCTGGCC 240

QY 241 CCGTGTACCTGTGCTGGGCCCCCGAGCTCCCGGTGGAGCTGTGAGTGTGGGCGGCA 300
DB 441 CCGTGTACCTGTGCTGGGCCCCCGAGCTCCCGGTGGAGCTGTGAGTGTGGGCGGCA 300
QY 301 GTCCAGTGGACCTGACAGATCCCTTTCATATGTGTCTCACTGTGGGCAATGATCTCAC 360
DB 501 GTCCAGTGGACCTGACAGATCCCTTTCATATGTGTCTCACTGTGGGCAATGATCTCAC 360
QY 361 GCGCTGTGACGCTTGCACATCATGAGCTGCATCTGCGCGGACCTTACATGGCAGC 420
DB 561 GCGCTGTGACGCTTGCACATCATGAGCTGCATCTGCGCGGACCTTACATGGCAGC 420
QY 421 GTGTACAAACGCGGACGCTGTGCGGCTTGTGTGGAGTGGGCGGCTGTGACCACTTC 480
DB 621 GTGTACAAACGCGGACGCTGTGCGGCTTGTGTGGAGTGGGCGGCTGTGACCACTTC 480
QY 481 TCGTCTGCTCTCTTCTACATCTGACAGCAGTGTCCACCGCGCTAGAGTGGCCAG 540
DB 681 TCGTCTGCTCTCTTCTACATCTGACAGCAGTGTCCACCGCGCTAGAGTGGCCAG 540
QY 541 ATGCAGACGACAGAGCTGCCAGCGCCACAGCTGTGTTCATGAGCTAGTGTGGCAGCA 600
DB 741 ATGCAGACGACAGAGCTGCCAGCGCCACAGCTGTGTTCATGAGCTAGTGTGGCAGCA 600
QY 601 CCGGACACCTCTAGCGGCTGGTGTACTCTCCCGGTCCGAGGAGGACACAGCCCTG 660
DB 801 CCGGACACCTCTAGCGGCTGGTGTACTCTCCCGGTCCGAGGAGGACACAGCCCTG 660
QY 661 GACCGGAGACAGCGGCGGCTGGAGCGCTTGGGACACAGGCTGTGTGGCCACCGTGTG 720
DB 861 GACCGGAGACAGCGGCGGCTGGAGCGCTTGGGACACAGGCTGTGTGGCCACCGTGTG 720
QY 721 AGCGAGTTGGGCTTGTGACGCGCACTATCTGATCTGCTGGGAGCAACGCTATCATC 780
DB 921 AGCGAGTTGGGCTTGTGACGCGCACTATCTGATCTGCTGGGAGCAACGCTATCATC 780
QY 781 TCGCGAGGAGAGCCCGTGGACGACACTACCTGGGGCTACTGCACTTTTGAAGATTTT 840
DB 981 TCGCGAGGAGAGCCCGTGGACGACACTACCTGGGGCTACTGCACTTTTGAAGATTTT 840
QY 841 TCCAAACTCTGGCTCTTCTCCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
DB 1041 TCCAAACTCTGGCTCTTCTCCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
QY 901 CAGAGCTTCCCGACGACAGCTCCCAACGCGTGATGAAAAAGCTGCCGCGGAGCCGAC 960
DB 1101 CAGAGCTTCCCGACGACAGCTCCCAACGCGTGATGAAAAAGCTGCCGCGGAGCCGAC 960
QY 961 TCGTCCCGGACCACTAGGGGGGTGACGAGCTGCTGGCGGTAG 1002
DB 1161 TCGTCCCGGACCACTAGGGGGGTGACGAGCTGCTGGCGGTAG 1002

RESULT 4
BC014241 1776 bp mRNA linear PRI 29-JUN-2004
LOCUS BC014241
DEFINITION Homo sapiens G protein-coupled receptor 146, mRNA (cdna clone
MGC:20737 IMAGE:456363), complete cds.
ACCESSION BC014241
VERSION BC014241.2 GI:40225566
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Klausner, R.D., Collins, P.S., Wagner, L., Shemen, C.M., Schlier, G.D.,
Altehoe, S.F., Zeeberg, B., Bueow, K.H., Schaefer, C.F., Bhac, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Cavasanti, T.L.,

Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toehiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loguelfano, N.A., Peters, G.J.,
Abramson, R.D., Mullin, S.J., Bosak, S.A., McEwen, P.J.,
McGernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Wotley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, D., Myers, R.M.,
Butterfield, J.S., Krzywiński, M.I., Skalska, U., Smalins, D.E.,
Schmerer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 1776)
Strausberg, R.
Direct Submission
Submitted (10-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Dec 19, 2003 this sequence version replaced gi:15559780.
Contact: MGC help desk
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hghri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larc, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,
Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 30 Row: J Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19923974.
location/Qualifiers
1..1776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20737 IMAGE:4563636"
/tissue_type="Kidney renal cell adenocarcinoma"
/clone_lib="NTH MGC_14"
/lab_host="DH10B-R"
/note="Vector: POTB7"
1..1776
/gene="GPR146"
/note="synonym: PGR8"
/db_xref="locusid:115330"
12..1013
/gene="GPR146"
/codon_start=1
/product="G protein-coupled receptor 146"
/protein_id="AAH14241.1"
/db_xref="GI:15559781"
/db_xref="locusid:115330"
/translation="MSSGSMFNGTGVEELPACODIQLGLSLSLGLVGVPGVGLCY
NALVLNHLNLSKSMTPDIVFVNMVAAGVLSALAFVHLILGPPSRMLMSVGEVH
VALQIPNVSLVAMYSTALISLIDHYTERALPRTYMASVYNTRHVCGFVWGALNLSF

ORIGIN
Query Match 100.0%; Score 1002; DB 9; Length 1776;
Best Local Similarity 100.0%; Pred. No. 3.4e-163;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGTGAGCTGAGCTGTTCAACCGCACAGGCTGTGTGAGAGCTGCTGCTGCCAG 60
12 ATGTGAGCTGAGCTGTTCAACCGCACAGGCTGTGTGAGAGCTGCTGCTGCCAG 71
61 GACTGCACTGGGCTGTCACTGTTGTGCTGCTGGGCTGTGTGAGAGCTGCTGCTG 120
72 GACTGCACTGGGCTGTCACTGTTGTGCTGCTGGGCTGTGTGAGAGCTGCTGCTG 131
121 GGCCTGTGCTAACAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
132 GGCCTGTGCTAACAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 191
181 ATGCCGAGCTGTACTTTGTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
192 ATGCCGAGCTGTACTTTGTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 251
241 CCGTGTGACCTGCTCGGCCCCCGAGCTCCCGTGTGAGAGAGAGAGAGAGAGAG 300
252 CCGTGTGACCTGCTCGGCCCCCGAGCTCCCGTGTGAGAGAGAGAGAGAGAGAG 311
301 GTCCAGCTGTGACCTGAGATCCCTTCAATGTGTCTCACTGTGTGCTGATCTCCAG 360
312 GTCCAGCTGTGACCTGAGATCCCTTCAATGTGTCTCACTGTGTGCTGATCTCCAG 371
361 GCCCTGTGAGCTGTGACCTGAGATCCCTTCAATGTGTCTCACTGTGTGCTGATCT 420
372 GCCCTGTGAGCTGTGACCTGAGATCCCTTCAATGTGTCTCACTGTGTGCTGATCT 431
421 GTGTACAAACCGGCGACCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
432 GTGTACAAACCGGCGACCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 491
481 TCCTGCTGCTCTTCAATCTGACAGCATGTGTGACAGAGAGAGAGAGAGAGAGAG 540
492 TCCTGCTGCTCTTCAATCTGACAGCATGTGTGACAGAGAGAGAGAGAGAGAGAG 551
541 ATGCAGAACGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
552 ATGCAGAACGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 611
601 CTGCGCACCTCTAGCGCGCTGTGTGCTAATCTCCCGCTCGGAGAGAGAGAGAG 660
612 CTGCGCACCTCTAGCGCGCTGTGTGCTAATCTCCCGCTCGGAGAGAGAGAGAG 671
661 GACCGGAGAACCGGCGCGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
672 GACCGGAGAACCGGCGCGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 731
721 AGCCAACTTTGGGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
732 AGCCAACTTTGGGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 791
781 TCGCGAGGAGAACCGGCGCGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
792 TCGCGAGGAGAACCGGCGCGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 851
841 TCGAAACTCTGGCTTCTTCCAGAGCTTGTGTGAGAGAGAGAGAGAGAGAGAG 900
852 TCGAAACTCTGGCTTCTTCCAGAGCTTGTGTGAGAGAGAGAGAGAGAGAGAG 911
901 CAGAGCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
912 CAGAGCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 971

QY	961	TGCTCCCGGACCATATGGGGGTGAGAGGTCGTGGGCTAG	1002
DB	972	TGCTCCCGGACCATATGGGGGTGAGAGGTCGTGGGCTAG	1013
RESULT 5			
LOCUS	AX230167	1778 bp	DNA
DEFINITION	Sequence 54 from Patent WO0162797.		
ACCESSION	AX230167		
VERSION	AX230167.1		GI:15592185
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Vogel, G., Wood, L.S., Parodi, L.A. and Lind, P.		
TITLE	Novel 9 protein-coupled receptors		
JOURNAL	Patent: WO 0162797-A 54 30-AUG-2001; PHARMACIA & UPJOHN COMPANY (US)		
FEATURES	Location/Qualifiers		
source	1..1778		
ORIGIN	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Query Match	100.0%;	Score 1002;	DB 6;
Best Local Similarity	100.0%;	Pred. No. 3.4e-153;	
Matches 1002;	Conservative	0;	Mismatches 0;
		Indels	0;
		Gaps	0;
QY	1	ATGTGAGACTGACGTGGTTCAACGGCAAGGAGGAGTGGAGAGAGCTGCTGGCCAG	60
DB	19	ATGTGAGACTGACGTGGTTCAACGGCAAGGAGTGGAGAGAGCTGCTGGCCAG	78
QY	61	GACCTGACGTGGGGCTGTCACTGTTGTGGCTGGGCTGTGGTGGGCGTGCAGTG	120
DB	79	GACCTGACGTGGGGCTGTCACTGTTGTGGCTGGGCTGTGGTGGGCGTGCAGTG	138
QY	121	GGCGTGTCTCAACAGCCCTGTGGTGGTGGGCAACCTACACAGAAAGCCAGATGACC	180
DB	139	GGCGTGTCTCAACAGCCCTGTGGTGGTGGGCAACCTACACAGAAAGCCAGATGACC	198
QY	181	ATGCGGACGTGTACTTGTCAACATGACAGTGGCAGGCTGTGCTCAGCGCCCTGGCC	240
DB	199	ATGCGGACGTGTACTTGTCAACATGACAGTGGCAGGCTGTGCTCAGCGCCCTGGCC	258
QY	241	CCTGTGCACCTGTGGCCCCCGAGCTCCGGTGGGCGCTGTGGAGTGTGGGCGGCGAA	300
DB	259	CCTGTGCACCTGTGGCCCCCGAGCTCCGGTGGGCGCTGTGGAGTGTGGGCGGCGAA	318
QY	301	GTCCACGAGGACATGACATCCCTCAATGTGTCTCACTGATGTGGGACATGTACCAAC	360
DB	319	GTCCACGAGGACATGACATCCCTCAATGTGTCTCACTGATGTGGGACATGTACCAAC	378
QY	361	GCCCTGTAGACTTGACCACTACATCGAGCGTGCATGCCGCGACCTACATGGCCAGC	420
DB	379	GCCCTGTAGACTTGACCACTACATCGAGCGTGCATGCCGCGACCTACATGGCCAGC	438
QY	421	GTTTACACACAGCGGCAAGTGTGGGCTTCTGTGTGGGTGGCGCGCTGCTTCAACCACTTC	480
DB	439	GTTTACACACAGCGGCAAGTGTGGGCTTCTGTGTGGGTGGCGCGCTGCTTCAACCACTTC	498
QY	481	TGCTGTGTCTCTTCTACATCTGACAGCATGTGTCAACCCGCGGCTAGATGGCCAG	540
DB	499	TGCTGTGTCTCTTCTACATCTGACAGCATGTGTCAACCCGCGGCTAGATGGCCAG	558
QY	541	ATGCAAAACGAGAAAGCTGCCAGCGCAAGCTGTGTTCAATGGCTACGTGTGCACGA	600
DB	559	ATGCAAAACGAGAAAGCTGCCAGCGCAAGCTGTGTTCAATGGCTACGTGTGCACGA	618
QY	601	CTGGGCAAGCTCTAGAGGCTGGTGTACTCTTCCGCGTCCGAGAGGAGGAACAGCGCCCTG	660

Db 619 CTGGCCACCCCTCTACGCGCTGTGTACTCTCCCGCTCCGCAAGGAGAGACACGCCCTTG 678

Qy 661 GACCGGAGCACACGGGCGGCTGGAGCCCTCGGCACACAGGCTGTGTGTGCGACCGTGTGC 720

Db 679 GACCGGAGCACACGGGCGGCTGGAGCCCTCGGCACACAGGCTGTGTGTGCGACCGTGTGC 738

Qy 721 ACGAGTTTGGGACTCTGGACGCGCACACTATCTGATCTCTGTGGGACACACGGTCATATC 780

Db 739 ACGAGTTTGGGACTCTGGACGCGCACACTATCTGATCTCTGTGGGACACACGGTCATATC 798

Qy 791 TCGCGAGGAAAGCCCGTGGACGCGCACACTACTGCGGGCTACTGCACTTTGTGAAGATTTTC 840

Db 799 TCGCGAGGAAAGCCCGTGGACGCGCACACTACTGCGGGCTACTGCACTTTGTGAAGATTTTC 858

Qy 841 TCCAAACTCTGGGCTTCTCCAGCAGCTTTGTGACACACTTCTCTACCGCTACATGAC 900

Db 859 TCCAAACTCTGGGCTTCTCCAGCAGCTTTGTGACACACTTCTCTACCGCTACATGAC 918

Qy 901 CAGAGCTTCCCAACCAAGCTTCCACACGGCTGATGAAAAAGCTGCCCTCCGGGACCGGCAC 960

Db 919 CAGAGCTTCCCAACCAAGCTTCCACACGGCTGATGAAAAAGCTGCCCTCCGGGACCGGCAC 978

Qy 961 TGCTTCCCGGACACATGCGGGTCCAGACGGTGTGGCGTAG 1002

Db 979 TGCTTCCCGGACACATGCGGGTCCAGACGGTGTGGCGTAG 1020

LOCUS	AX549394	1785 bp	DNA	linear	PAT 26-NOV-2002
DEFINITION	Sequence 679 from Patent WO02061087.				
ACCESSION	AX549394				
VERSION	AX549394.1 GI:25814007				
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Buxarjofa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1				
AUTHORS	Butner,G.C., Roush,C.L. and Brown,J.P.				
TITLE	Antigenic peptides, such as for G protein-coupled receptors (GPCRs), antibodies thereto, and systems for identifying such antigenic peptides				
JOURNAL	Patent: WO 02061087-A 679 08-AUG-2002; Lifespan Biosciences, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..1785				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	100.0%; Score 1002; DB 6; Length 1785;				
Best Local Similarity	100.0%; Prd. No.3.4e-163;				
Matches 1002; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	ATGTGAGAGCTGACAGCTGGTTCAACGGCAGAGGCTGTGTGAGAGAGCTGCTGCGCAG	60		
DB	21	ATGTGAGAGCTGACAGCTGGTTCAACGGCAGAGGCTGTGTGAGAGAGCTGCTGCGCAG	80		
QY	61	GACCTGCAGAGCTGGGGCTGTCACTGTTGTCGCTGCTGGGCTGTGTGTGGGCTGTCACAGT	120		
DB	81	GACCTGCAGAGCTGGGGCTGTCACTGTTGTCGCTGCTGGGCTGTGTGTGGGCTGTCACAGT	140		
QY	121	GAGCTGTGCTCAACAGCGCCCTGCTGGTGTGCGCAACTACACAGACAGGCGCAGATAGC	180		
DB	141	GAGCTGTGCTCAACAGCGCCCTGCTGGTGTGCGCAACTACACAGACAGGCGCAGATAGC	200		
QY	181	ATGCGGAGCTGTACTTTGTCAACATGAGCAGTGGCAGGCTGTGTGTCAAGCGCCCTGACC	240		
DB	201	ATGCGGAGCTGTACTTTGTCAACATGAGTGGCAGGCTGTGTGTCAAGCGCCCTGACC	260		

```

Oy 241 CCTGTGACCTGCTGGGCCCCCGAGCTCCGGTGGGCGCTGTGGAGTGTGGGCGGAGAA 300
Db 261 CCTGTGACCTGCTGGGCCCCCGAGCTCCGGTGGGCGCTGTGGAGTGTGGGCGGAGAA 320
Oy 301 GTCCACGTGGGACATGCAATCCCTTCAATGTCTCACTGTGGGCGATGACTCCAGC 360
Db 321 GTCCACGTGGGACATGCAATCCCTTCAATGTCTCACTGTGGGCGATGACTCCAGC 380
Oy 361 GCGCTGTGACCTGACACCACTACATGAGGCTGACATGCGCGGAGACCTACATAGGCGAC 420
Db 381 GCGCTGTGACCTGACACCACTACATGAGGCTGACATGCGCGGAGACCTACATAGGCGAC 440
Oy 421 GTTACAAACAGCGGCGACGTGTGGCTGTGTGGGAGTGGCGCGCTGACCAAGCTTC 480
Db 441 GTTACAAACAGCGGCGACGTGTGGCTGTGTGGGAGTGGCGCGCTGACCAAGCTTC 500
Oy 481 TCCTCGCTGCTCTTCTACATCTGACGCAATGTCCACCCCGCGCTGACAGTGGCGCAAG 540
Db 501 TCCTCGCTGCTCTTCTACATCTGACGCAATGTCCACCCCGCGCTGACAGTGGCGCAAG 560
Oy 541 ATGCAAGACGAGAGCTGCGACGCGACCGCTGGTTCATCGGCTACGTTGTCAGGCA 600
Db 561 ATGCAAGACGAGAGCTGCGACGCGACCGCTGGTTCATCGGCTACGTTGTCAGGCA 620
Oy 601 CTGGCGACCTCTTACGCGCTGTGTCTACTCTCCGCGTCCGACGAGGAGACAGCCCTTG 660
Db 621 CTGGCGACCTCTTACGCGCTGTGTCTACTCTCCGCGTCCGACGAGGAGACAGCCCTTG 680
Oy 661 GACCGGAGACAGGCGCGCTGTGAGCCCTGCGGACACAGGCTGTGTGGCGACCGCTTGC 720
Db 681 GACCGGAGACAGGCGCGCTGTGAGCCCTGCGGACACAGGCTGTGTGGCGACCGCTTGC 740
Oy 721 ACCGAGTTTGGGCTCTGAGCGCACTATCTGATCTGCTGGGCGACAGGCTATCATC 780
Db 741 ACCGAGTTTGGGCTCTGAGCGCACTATCTGATCTGCTGGGCGACAGGCTATCATC 800
Oy 781 TCSCGAGGAGAGCCCGTGGACGACACTACCTGGGCTACTGCACTTTGTGAAGATTTC 840
Db 801 TCSCGAGGAGAGCCCGTGGACGACACTACCTGGGCTACTGCACTTTGTGAAGATTTC 860
Oy 841 TCCAAACTCTGGCTCTTCCAGCAGCTTTGTGACACCACTCTTACCGCTACATGAAC 900
Db 861 TCCAAACTCTGGCTCTTCCAGCAGCTTTGTGACACCACTCTTACCGCTACATGAAC 920
Oy 901 CAGAGTTTCCCGAGCAAGCTCCACGCGCTGATGAAAAAGCTGCGCGGAGACCGGAC 960
Db 921 CAGAGTTTCCCGAGCAAGCTCCACGCGCTGATGAAAAAGCTGCGCGGAGACCGGAC 980
Oy 961 TGCCTCCCGGACCAATGGGGGTGACAGAGGTGCTGGCGTAG 1002
Db 981 TGCCTCCCGGACCAATGGGGGTGACAGAGGTGCTGGCGTAG 1022

```

```

RESULT 7
AC140819 188370 bp DNA linear HTG 03-MAR-2003
LOCUS Homo sapiens chromosome 5 clone RP11-734E9, WORKING DRAFT SEQUENCE,
AC140819 8 unordered pieces.
ACCESSION AC140819
VERSION AC140819.1 GI:28631200
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 188370)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 188370)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2003) Production Sequencing Facility, DOE Joint

```

COMMENT

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

 Project Information
 Center Project Name: 1576465
 Center clone name: RP11-11_734E9

Summary Statistics

Consensus quality: 183379 bases at least Q40
 Consensus quality: 184275 bases at least Q30
 Consensus quality: 184958 bases at least Q20
 Estimated insert size: 175000; agarose-fp estimation
 Estimated insert size: 187670; sum-of-contigs estimation
 Quality coverage: 9.36 in Q20 bases; agarose-fp estimation
 Quality coverage: 8.73 in Q20 bases; sum-of-contigs estimation.
 NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 1155: contig of 1155 bp in length
* 1156 1255: gap of unknown length
* 1256 4105: contig of 2850 bp in length
* 4106 4205: gap of unknown length
* 4206 10555: contig of 6350 bp in length
* 10556 10655: gap of unknown length
* 10656 24895: contig of 14240 bp in length
* 24896 24995: gap of unknown length
* 24996 42191: contig of 17196 bp in length
* 42192 42291: gap of unknown length
* 42292 60204: contig of 17913 bp in length
* 60205 60304: gap of unknown length
* 60305 112109: contig of 51805 bp in length
* 112110 112209: gap of unknown length
* 112210 188370: contig of 76161 bp in length.

```

FEATURES

```

source
1..188370
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-734E9"
/location="RP11 human BAC library 11"

```

ORIGIN

Query Match 100.0%; Score 1002; DB 2; Length 188370;
 Best Local Similarity 100.0%; Pred. No. 2e-163;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 ATGTGAGCTGACAGCTGTTCAACGACGACAGGCTGTGAGAGAGTGCCTGCCAG 60
Db 46995 ATGTGAGCTGACAGCTGTTCAACGACGACAGGCTGTGAGAGAGTGCCTGCCAG 47054
Oy 61 GACCTGACCTGGGGGTGTCACTGTTGTGCTGCTGGGCGTGTGTGGCGCTGCAAGT 120
Db 47055 GACCTGACCTGGGGGTGTCACTGTTGTGCTGCTGGGCGTGTGTGGCGCTGCAAGT 47114
Oy 121 GGCCTGTGTACACAGCCCTGTGGTGTGGGCACTACACAGAGGACGACTGAC 180
Db 47115 GGCCTGTGTACACAGCCCTGTGGTGTGGGCACTACACAGAGGACGACTGAC 47174
Oy 181 ATGCGGACGTGTACTTTGTCAATGACATGACAGTGGAGGCTGTGCTGAGGCTGAGCC 240
Db 47175 ATGCGGACGTGTACTTTGTCAATGACATGACAGTGGAGGCTGTGCTGAGGCTGAGCC 47234
Oy 241 CCTGTGACCTGCTGGGCCCCCGAGCTCCGGTGGGCGCTGTGAGTGTGGGCGGAGAA 300

```

Db	47235	CTGTGACACCTGCTGGGCCCCCGAGCTCCCGAGTGGGCGGTGTGAGATGTGAGGCGGAGAA	47299
Qy	301	GTCCACGTGGGCACTGACAGATCCCTTTCAATGTGTCTCTACCTGGGCACTGATCTCACC	360
Db	47295	GTCCACGTGGGCACTGACAGATCCCTTTCAATGTGTCTCTACCTGGGCACTGATCTCACC	47354
Qy	361	GGCCTGTGAGCCCTCGACCACTACATGAGGGTGCATGTGCGGGGACCTTACATGGGCAAG	420
Db	47355	GGCCTGTGAGCCCTCGACCACTACATGAGGGTGCATGTGCGGGGACCTTACATGGGCAAG	47411
Qy	421	GTTGACAAACAGCGGCAAGTGTGGGCTTGTGTGGGGTGGGCGGCTGTACCAAGCTTC	480
Db	47415	GTTGACAAACAGCGGCAAGTGTGGGCTTGTGTGGGGTGGGCGGCTGTACCAAGCTTC	47477
Qy	481	TCTCTGCTGCTCTTCTACATCTGTGACCAATGTGTCCACCCGCGGCTAGAGTGGCCAA	540
Db	47475	TCTCTGCTGCTCTTCTACATCTGTGACCAATGTGTCCACCCGCGGCTAGAGTGGCCAA	47533
Qy	541	ATGAGAAACGAGAAAGCTGGCGAGAGCGCAAGCTGTGTCTTACCTGCGTACGTGTGGCCAG	600
Db	47535	ATGAGAAACGAGAAAGCTGGCGAGAGCGCAAGCTGTGTCTTACCTGCGTACGTGTGGCCAG	47599
Qy	601	CTGGCCACCCCTCTACGCGCTGTGTCTACTCTCCGCGCTCCGAGAGAGACAGCCCTTG	660
Db	47595	CTGGCCACCCCTCTACGCGCTGTGTCTACTCTCCGCGCTCCGAGAGAGACAGCCCTTG	47655
Qy	661	GACCGGGACACAGGGGCGGCTGTGAGACCTCTCGGACACAGGGTGTGTGTGGCACTGTGTC	720
Db	47655	GACCGGGACACAGGGGCGGCTGTGAGACCTCTCGGACACAGGGTGTGTGTGGCACTGTGTC	47711
Qy	721	ACGCAATTTTGGGCTCTGGAAGGCCACATATCTGTACTCTGTGGGGGACAGGTCATATTC	780
Db	47715	ACGCAATTTTGGGCTCTGGAAGGCCACATATCTGTACTCTGTGGGGGACAGGTCATATTC	47771
Qy	781	TCCGAGAGAAAGCCCGTGAAGGACACATACCTGGGGCTACCTGCACTTTGAGAGATTTC	840
Db	47775	TCCGAGAGAGAAAGCCCGTGAAGGACACATACCTGGGGCTACCTGCACTTTGAGAGATTTC	47833
Qy	841	TCCAAATCTCTGGGCTTCTCCAGAGGTTTGTGCACACACTTCTCTACCGCTACATGAAC	900
Db	47835	TCCAAATCTCTGGGCTTCTCCAGAGGTTTGTGCACACACTTCTCTACCGCTACATGAAC	47899
Qy	901	CAGAGCTTCCCGACCAAGCTTCCAAAGGCTGATGAAAAAGCTGGCTGCGGGGACCGGCAAC	960
Db	47895	CAGAGCTTCCCGACCAAGCTTCCAAAGGCTGATGAAAAAGCTGGCTGCGGGGACCGGCAAC	47959
Qy	961	TGCTCCCCGAGCAATGGGGGGTGCAGCAGGTGCTGGCGTAG	1002
Db	47955	TGCTCCCCGAGCAATGGGGGGTGCAGCAGGTGCTGGCGTAG	47996
RESULT 8			
LOCUS	AC073957	196204 bp	DNA linear PRI 09-JAN-2002
DEFINITION	Homo sapiens BAC clone Rpl1-449p15 from 7, complete sequence.		
ACCESSION	AC073957		
VERSION	AC073957.7	GI:17149791	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 196204)		
AUTHORS	Sulston,J.B. and Waterston,R.		
TITLE	Toward a complete human genome sequence		
JOURNAL	Genome Res. 8 (11), 1097-1108 (1998)		
MEDLINE	99063792		
PUBMED	9847074		
REFERENCES	2 (bases 1 to 196204)		
AUTHORS	Leonard,S., Cotton,M., Hawkins,M. and Spalding,L.		
TITLE	The sequence of Homo sapiens BAC clone Rpl1-449p15		
JOURNAL	Unpublished (2001)		
REFERENCE	3 (bases 1 to 196204)		

REFERENCE	REFERENCE
AUTHORS	AUTHORS
TITLE	TITLE
JOURNAL	JOURNAL
COMMENT	

Waterston, R.H.
Direct Submission
Submitted (07-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 196204)
Waterston, R.H.
Direct Submission
Submitted (29-NOV-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 196204)
Waterston, R.H.
Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 196204)
Waterston, R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 29, 2001 this sequence version replaced gt:15589737.

Center code: WTGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics -----
Center project name: H_MN0449P15

Center project name: H_NH0449P15

9P15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GR7/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCL1 human BAC library was made from the blood of one male donor, as described by Oosagawa, K., Moon, P. Y., Zhao, B., Frangen, E., Tateo, M., Catanesi, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.regen.com>) or Piter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-1244M4; the clone sequenced to the right is RP13-580113. Actual start of this clone is at base position 1 of RP11-449P15; actual end is at base position 196204 of RP11-449P15.

Data from AC091729 was used to finish this clone, AC073957. Polymorphisms have been identified between AC073957 and AC091729.

Db	1790556	TCCTGGCTGCTCTTCTTACATCTGACAGCCANTGTGTCCACC	CGCGCTAGAGTGTCCGACAA	179115
QY	541	ATGCAAAACGCAAAAGCTGCTCCGACGCAAGCTGTGTTCAT	TGAGCTACGTGTGTCCAGCA	600
Db	179116	ATGCAAAACGCAAAAGCTGCTCCAGCGCAAGCTGTGTTCAT	TGAGCTACGTGTGTCCAGCA	179175
QY	601	CTGGACACCCCTTACGCGCTGTGNGCTACTCTCCGCGGTCCG	CAGGAGGACACAGCCCTG	660
Db	179176	CTGGACACCCCTTACGCGCTGTGNGCTACTCTCCGCGGTCCG	CAGGAGGAGACAGCCCTG	179235
QY	661	GAACGGGACACAGGCGCGCTGTGAGCCCTTCGACACAGGCT	GCTGTGACCAAGCTGTGC	720
Db	179236	GAACCGGACACAGGCGCGCTGTGAGCCCTTCGACACAGGCT	GCTGTGTGACCAAGCTGTGC	179295
QY	721	ACGCAAGTTTGGGCTGTGGACGCGCACATATCTGATCTCTG	CTGGGGGACACAGGCTATCATC	780
Db	179296	ACGCAAGTTTGGGCTGTGGACGCGCACATATCTGATCTCTG	CTGGGGGACACAGGCTATCATC	179355
QY	781	TCGCGAGGGAGAGCCCGGTGACGCACTTACCTTGGGGCTA	CTGCACTTTGTGAAGATTT	840
Db	179356	TCGCGAGGGAGAGCCCGGTGACGCACTTACCTTGGGGCTA	CTGCACTTTGTGAAGATTT	179415
QY	841	TCCAAACTCTGCGCCCTTCTCCAGCAGCTTTGTGACACCA	CTTCTTACACGCTACATGAAC	900
Db	179416	TCCAAACTCTGCGCCCTTCTCCAGCAGCTTTGTGACACCA	CTTCTTACACGCTACATGAAC	179475
QY	901	CAGAGCTTCCCGACAGAGCTCCCAAGCGCTGATGAAGAAAC	CTGCGCTGCGGGACCGGCA	960
Db	179476	CAGAGCTTCCCGACAGAGCTCCCAAGCGCGTGAAGAAAGAA	CTGCGCTGCGGGAGCGGCA	179535
QY	961	TGCTTCCCGGACACATGAGGGGTGACGACGGTGTGCGCTAG	1002	
Db	179536	TGCTTCCCGGACACATGAGGGGTGACGACGGTGTGCGCTAG	179577	

RESULT 9	AR242245	1614 bp	DNA	linear	PAT 20-DEC-2002
LOCUS	AR242245				
DEFINITION	Sequence 2 from patent US 6472173.				
ACCESSION	AR242245				
VERSION	AR242245.1	GI:27288068			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1614)				
TITLE	Ford, J. and Yeung, G.				
JOURNAL	Chemokine receptor obtained from a cDNA library of fetal				
FEATURES	liver-spleen				
source	Patent: US 6472173-A 2 29-OCT-2002;				
	Location/Qualifiers				
	1..1614				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					

	Query Match	Similarity	%	Score	1000 / 8 :	DB	6 :	Length	1614 ;	
	Best Local	Similarity	99.7%	Pred.	No. 5-5e-163;					
	Matches	999 ;	Conservative	3 ;	Mismatches	0 ;	Indels	0 ;	Gaps	0
Oy	1	ATGTGGAGCTGCAGCTGTTCAACGGACA	CAGGGCTGGTGAAGAGCTGCCCTTGCCAG	60	:	:	:	:	:	:
Dd	129	ATGTGAACTGACACTGGTTTCAA	CGGCACARGGCTGGTGAAGAGCTGTCCTCGCCAR	188	:	:	:	:	:	:
Oy	61	GACCTGCAGCTGGGGCTGTCACTGTTTGTGCTGCTGGGCTTGGTGGAGCTGECACATG	120	:	:	:	:	:	:	:
Dd	189	GACCTGCAGCTGGGGCTGTCACTGTTTGTGCTGCTGGGCTTGGTGGAGCTGECACATG	248	:	:	:	:	:	:	:
Oy	121	GACCCTGCTCTAACAGGCCCTGCTGGTGGTGGCCCACTTCA	CAGCAAGGCAGCATTTGACC	180	:	:	:	:	:	:
Dd	249	GACCCTGCTCTAACAGGCCCTGCTGGTGGTGGCCCACTTCA	CAGCAAGGCAGCATTTGACC	308	:	:	:	:	:	:
Oy	181	ATGCCGAGCAGTACTTTTGTCAACATGGCAGTGGCAGGCTGATGCTCAGCCGCTCTGGCC	240	:	:	:	:	:	:	:

Db	309	ATGCGGAGCGGTACCTTTGTCAACATGGCAGTGGCAAGCCCTGGTCTCAGGCGCCTGGCC	368
Qy	241	CTGTGCACTGTCTCGGCGCCCCCGAGCTCCCGGTGGCGCTGTGAAGTGTGG-CGGCGAA	300
Db	369	CCTGTGACCTGTCTCGGCGCCCCCGAGCTCCCGGTGGCGCTGTGAAGTGTGGCGGCGAA	428
Qy	301	GTCCACGTGGACCTGCAAGATCCCTCTCAAATGTCTCTCACTGGTGGCCATGTACTTCACC	360
Db	429	GTCCACGTGGACCTGCAAGATCCCTCTCAAATGTCTCTCACTGGTGGCCATGTACTTCACC	488
Qy	361	GCCGTGCTGAGCCTCGACCACTACATCGAGGGTGTCACTGGCGGGAACCTACATGAGCAGC	420
Db	489	GCCCTGCTGAGCCTCGACCACTACATGAGGGTGTCACTGGCGGGAACCTACATGAGCAGC	548
Qy	421	GTGTACAAACAAGCGGACAGTGTGGGCTTCGTGTGGGGTGGCGGCGCTGTGACCAAGCTTC	480
Db	549	GTGTACAAACAAGCGGACAGTGTGGGCTTCGTGTGGGGTGGCGGCGCTGTGACCAAGCTTC	608
Qy	481	TCTCTGCTGCTCTTCTACATGTGACGCAATGTGTCCACCGCGCGGCTTAAATGTGGCCAAAG	540
Db	609	TCTCTGCTGCTCTTCTACATGTGACGCAATGTGTCCACCGCGCGGCTTAAATGTGGCCAAAG	668
Qy	541	ATGAGAAACGAGAAAGCTGCCAGCGGCAAGCTGGTGTTCATCGGCTCTAGTGTGCACGA	600
Db	669	ATGAGAAACGAGAAAGCTGCCAGCGGCAAGCTGGTGTTCATCGGCTCTAGTGTGCACGA	728
Qy	601	CTGGCCACCCCTCTACGCGCTGTGTCTACTCTCCCGCTCCGACAGGAGAGAACCCCTG	660
Db	729	CTGGCCACCCCTCTACGCGCTGTGTCTACTCTCCCGCTCCGACAGGAGAGAACCCCTG	788
Qy	661	GACCGGAGCAACGGGCGGCTGGAGCCCTCGGACACAGGCTGTGTGGTGGCAACCGTGTGC	720
Db	789	GACCGGAGCAACGGGCGGCTGGAGCCCTCGGACACAGGCTGTGTGGTGGCAACCGTGTGC	848
Qy	721	ACGCAATTGGGCTCTGTGAGCGCCCACTATCTGTATCTCTGTGTGGGACACAGGTATCATTC	780
Db	849	ACGCAATTGGGCTCTGTGAGCGCCCACTATCTGTATCTCTGTGTGGGACACAGGTATCATTC	908
Qy	781	TGCGAGAGGAAGCCCGTGGAGCGCACACTACCTGGGGGTACATGCACTTTGTGTAAGATTTTC	840
Db	909	TGCGAGAGGAAGCCCGTGGAGCGCACACTACCTGGGGGTACATGCACTTTGTGTAAGATTTTC	968
Qy	841	TCCAAACTCTGAGCCTTCTCCAGAGCTTTGTGTGACACACTTCTCTACCCGTACATGAAC	900
Db	969	TCCAAACTCTGAGCCTTCTCCAGAGCTTTGTGTGACACACTTCTCTACCCGTACATGAAC	1028
Qy	901	CAGAGCTTCCCGACGAAGTTCAAAGGCTGTGAAAAAGCTGCGCTGTGGGGAACCGGAC	960
Db	1029	CAGAGCTTCCCGACGAAGTTCAAAGGCTGTGAAAAAGCTGCGCTGTGGGGAACCGGAC	1088
Qy	961	TGCTTCCCCGAGCAATGGGGGTGGACAAGTGTGGCGTAG	1002
Db	1089	TGCTTCCCCGAGCAATGGGGGTGGACAAGTGTGGCGTAG	1130

RESULT 10				
AX230144	AX230144	1750 bp	DNA	linear
LOCUS	AX230144	Sequence 31 from Patent WO0162797.		
DEFINITION	AX230144			
ACCESSION	AX230144			
VERSION	AX230144.1	GI:15592165		
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1			
AUTHORS	Vogel, G., Wood, L.S., Parodi, L.A., and Lind, P.			
TITLE	Novel G protein-coupled receptors			
JOURNAL	Patent: WO 0162797-A 31 30 AUG-2001;			
	PHARMACIA & UPJOHN COMPANY (US)			
FEATURES	Location/Qualifiers			

source 1..1750
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

```
Query Match 99.8%; Score 1000.4; DB 6; Length 1750;
Best Local Similarity 99.9%; Pred. No. 6,4e-163;
Matches 1001; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGCAGCTGCTTCAACGGCAAGGGCTGTGAGGAGCTGCTGCTCCAG 60
DB 139 ATGTGAGCTGCAGCTGCTTCAACGGCAAGGGCTGTGAGGAGCTGCTGCTCCAG 198
QY 61 GACCTCAGCTGGGGCTGTCACTGTGTGCTGCTGCTGGGCTGGTGGGCGGCGCAAG 120
DB 199 GACCTCAGCTGGGGCTGTCACTGTGTGCTGCTGCTGGGCTGGTGGGCGGCGCAAG 258
QY 121 GGGCTGTGCTACAAAGCCCTGCTGTGCTGGCCCACTACAGCAAGCCAGCATAGCC 180
DB 259 GGGCTGTGCTACAAAGCCCTGCTGTGCTGGCCCACTACAGCAAGCCAGCATAGCC 318
QY 181 ATGCCGGAAGTGAATCTTTGTCAACATGGCAATGGCAAGGCTGTGCTAGCGGCTGCGC 240
DB 319 ATGCCGGAAGTGAATCTTTGTCAACATGGCAAGGCTGTGCTAGCGGCTGCGC 378
QY 241 CCTGTGCACTGCTGCTGGCCCGGCAAGCTCCGGTGGGCGCTGTGAGAGTGTGGGCGGCA 300
DB 379 CCTGTGCACTGCTGCTGGCCCGGCAAGCTCCGGTGGGCGCTGTGAGAGTGTGGGCGGCA 438
QY 301 GTTCACGTGGCACTGCAAGATCCCTTCAATGTGTCTCACTGTGGCCATGTACTCACC 360
DB 439 GTTCACGTGGCACTGCAAGATCCCTTCAATGTGTCTCACTGTGGCCATGTACTCACC 498
QY 361 GGCCTGCTAGCTGCAACCACTACATGAGCTGCACTCCCGGCACTCACTAGCCAGC 420
DB 499 GGCCTGCTAGCTGCAACCACTACATGAGCTGCACTCCCGGCACTCACTAGCCAGC 558
QY 421 GTGTACAAACAGCGGCACTGTGGGCTTGTGTGGGAGTGGGCGCTGTGCAAGCTTC 480
DB 559 GTGTACAAACAGCGGCACTGTGGGCTTGTGTGGGAGTGGGCGCTGTGCAAGCTTC 618
QY 481 TCTGTGCTCTTTCTTCACTGTGCAAGCATGTGTCCACCCGCGCTAGAGTGCACAG 540
DB 619 TCTGTGCTCTTTCTTCACTGTGCAAGCATGTGTCCACCCGCGCTAGAGTGCACAG 678
QY 541 ATGCAAGACGGAAGCTGCCGCAACGCAAGCTGTGTTATGCTGCTAGCTGTGCCAGCA 600
DB 679 ATGCAAGACGGAAGCTGCCGCAACGCAAGCTGTGTTATGCTGCTAGCTGTGCCAGCA 738
QY 601 CTGGCCACCTCTTACGCGCTGTGCTACTCTCCGCGTCCGAGGAGGAGCAAGCCCTG 660
DB 739 CTGGCCACCTCTTACGCGCTGTGCTACTCTCCGCGTCCGAGGAGGAGCAAGCCCTG 798
QY 661 GACCGGAGACGGGCGGCTGAGACCTTGGGACACAGGCTGTGTGGCCACCGTGTGC 720
DB 799 GACCGGAGACGGGCGGCTGAGACCTTGGGACACAGGCTGTGTGGCCACCGTGTGC 858
QY 721 AGCGAGTTTGGGCTTGAACGCGCACTATCTGATCTGCTGGGCAACAGCTCATCATC 780
DB 859 AGCGAGTTTGGGCTTGAACGCGCACTATCTGATCTGCTGGGCAACAGCTCATCATC 918
QY 781 TCGCAGAGGAGAGCCGCTGAGCAGCACTAAGCTGGGCTACTGACTTTGTGAAGATTTC 840
DB 919 TCGCAGAGGAGAGCCGCTGAGCAGCACTAAGCTGGGCTACTGACTTTGTGAAGATTTC 978
QY 841 TCGAAACTCTGGCTTTCTTCCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAA 900
DB 979 TCGAAACTCTGGCTTTCTTCCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAA 1038
QY 901 CAGAGCTTCCCGAGAGCTCCAAAGGCTGATGAAAAGCTGCTGTGGGGAGCCGGGAC 960
DB 1039 CAGAGCTTCCCGAGAGCTCCAAAGGCTGATGAAAAGCTGCTGTGGGGAGCCGGGAC 1098
```

```
QY 961 TGCTCCCGGACCAATAGGGGTGCAAGAGGCTGCGGCTAG 1002
DB 1099 TGCTCCCGGACCAATAGGGGTGCAAGAGGCTGCGGCTAG 1140
```

```
RESULT 11
BD231040
LOCUS BD231040 1878 bp DNA linear PAT 17-JUL-2003
DEFINITION Human GPCR proteins.
ACCESSION BD231040
VERSION BD231040.1 GI:33040810
KEYWORDS JP 2002525054-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1878)
Bandman,O., Lal,P., Tang,T.Y., Corley,N.C., Guegler,K.J.,
Gorgone,G.A. and Baughn,M.R.
Human GPCR proteins
Patent: JP 2002525054-A 3 13-AUG-2002;
INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002525054-A/3
PD 13-AUG-2002
PF 17-SEP-1999 JP 2000570320
PR 17-SEP-1998 US 09/156513
PI OLGA BANDMAN,PREETI LAL,TOM Y TANG,NEIL C CORLEY,KARL J PI
GUEGLER,
PI GINA A GORGONE,MARIAH R BAUSHN
PC C12N15/09,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61P1/04,
A61P1/16,
PC
A61P1/18,A61P3/00,A61P3/10,A61P5/16,A61P7/00,A61P7/04,A61P7/06,PC
A61P9/10,
PC A61P9/10,A61P11/00,A61P11/06,A61P13/02,A61P13/12,A61P17/00,PC
A61P17/02,
PC A61P17/06,A61P17/12,A61P19/00,A61P19/02,A61P19/06,A61P19/08,
PC A61P19/10,
PC A61P21/02,A61P21/04,A61P25/00,A61P25/02,A61P25/08,A61P25/14,
PC A61P25/16,
PC A61P25/18,A61P25/28,A61P27/02,A61P29/00,A61P29/00,A61P31/00,
PC A61P31/04,
PC A61P31/10,A61P31/12,A61P31/18,A61P35/00,A61P35/02,A61P37/00,
PC A61P37/02,
PC A61P37/08,C07K14/705,C12N1/15,C12N1/19,C12N1/21,C12N5/10,PC
C12P21/02,
PC C12Q1/68,C12N15/00,C12N5/00,A61K37/02
CC Incyte ID No: 2214673CB1
FH Key
FT source 1..1878
location/Qualifiers
1..1878
/organism="Homo sapiens (human)".
```

FEATURES

source

location/Qualifiers
1..1878
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

```
Query Match 99.8%; Score 1000.4; DB 6; Length 1878;
Best Local Similarity 99.9%; Pred. No. 6,4e-163;
Matches 1001; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGCAGCTGCTTCAACGGCAAGGGCTGTGAGGAGCTGCTGCTGCCAG 60
DB 139 ATGTGAGCTGCAGCTGCTTCAACGGCAAGGGCTGTGAGGAGCTGCTGCTGCCAG 198
QY 61 GACCTCAGCTGGGGCTGTCACTGTGTGCTGCTGGGCTGGTGGGCGGCGCAAG 120
DB 199 GACCTCAGCTGGGGCTGTCACTGTGTGCTGCTGGGCTGGTGGGCGGCGCAAG 258
QY 121 GGGCTGTGCTACAAAGCCCTGCTGTGCTGGCCCACTACAGCAAGCCAGCATAGCC 180
```

```

Db      259 GGGCTGTGTCACAAAGCCCTGCTGGTGGCCAACTACAGCAAGGCGACGATGACC 318
      181 ATGCCGAGAGTGTACTTTGTCAACATGGCAGTGGCGGCTGTCTCAGCGCCCTGGCC 240
      319 ATGCCGAGAGTGTACTTTGTCAACATGGCAGTGGCGGCTGTCTCAGCGCCCTGGCC 378
      241 CCTGTGACCTGTGCTGGGCCCCGAGAGCTCCCGGTGGGCGCTGTGGAGTGTGGGCGGAA 300
      379 CCTGTGACCTGTGCTGGGCCCCGAGAGCTCCCGGTGGGCGCTGTGGAGTGTGGGCGGAA 438
      301 GTCCAGTGGCAGTGTGAGATCCCTCTCAATGTGTCTCACTGTGTGGCCATGTACTCCAC 360
      439 GTCCAGTGGCAGTGTGAGATCCCTCTCAATGTGTCTCACTGTGTGGCCATGTACTCCAC 498
      361 GCCCTGTGAGCTGTGAGCACTATGATGAGCGTGTGCTGCGCGGACCTATCATGCCACG 420
      499 GCCCTGTGAGCTGTGAGCACTATGATGAGCGTGTGCTGCGCGGACCTATCATGCCACG 558
      421 GTGTACAAACAGCGGAGCGGTGTGGGCTGTGGTGGGCGTGTGGTGTGACCACTTC 480
      559 GTGTACAAACAGCGGAGCGGTGTGGGCTGTGGTGGGCGTGTGGTGTGACCACTTC 618
      481 TCTCGTGTGTCTTTTCAATCTGTGACGCAATGTGTCCACCGCGCGCTAGAGTGGCGCAAG 540
      619 TCTCGTGTGTCTTTTCAATCTGTGACGCAATGTGTCCACCGCGCGCTAGAGTGGCGCAAG 678
      541 ATGCAGAACGCAAGAGTGTCCGACGCAAGCGTGTGTATCGGCTACTGTGTGCCAGCA 600
      679 ATGCAGAACGCAAGAGTGTCCGACGCAAGCGTGTGTATCGGCTACTGTGTGCCAGCA 738
      601 CTGGCCACCTCTACGCGGTGTGTGTACTCTCCCGGTGCGGAGGAGGACACGCCCTG 660
      739 CTGGCCACCTCTACGCGGTGTGTGTACTCTCCCGGTGCGGAGGAGGACACGCCCTG 798
      661 GACCGGAGACAGCGGCGCGCTGTGAGCGCTGTGAGCGCAAGGCTGTGTGGCCACCTGTGTC 720
      799 GACCGGAGACAGCGGCGCGCTGTGAGCGCTGTGAGCGCAAGGCTGTGTGGCCACCTGTGTC 858
      721 AGCGAGTTGGGCTGTGAGCGCACTATGTGATCTGTGTGGGCGCAACGGTCTATCATC 780
      859 AGCGAGTTGGGCTGTGAGCGCACTATGTGATCTGTGTGGGCGCAACGGTCTATCATC 918
      781 TGGCGAGGAGAGCGCGGTGTGAGCGCACTATGTGATCTGTGTGGGCGCAACGGTCTATCATC 840
      919 TGGCGAGGAGAGCGCGGTGTGAGCGCACTATGTGATCTGTGTGGGCGCAACGGTCTATCATC 978
      841 TCCAAACTCTGGCTGTCTCAGCAGCTTGTGTGACCACTTCTTACCGCTACATGAGC 900
      979 TCCAAACTCTGGCTGTCTCAGCAGCTTGTGTGACCACTTCTTACCGCTACATGAGC 1038
      901 CAGAGCTTCCCGAGCAAGCTTCCAAAGCGGTGTGATGAAAAAGCTGTGGCGGAGCGGCAAC 960
      1039 CAGAGCTTCCCGAGCAAGCTTCCAAAGCGGTGTGATGAAAAAGCTGTGGCGGAGCGGCAAC 1098
      961 TGTCTCCCGAGCAAGCTTCCAAAGCGGTGTGATGAAAAAGCTGTGGCGGAGCGGCAAC 1002
      1099 TGTCTCCCGAGCAAGCTTCCAAAGCGGTGTGATGAAAAAGCTGTGGCGGAGCGGCAAC 1140

```

```

REFERENCE
AUTHORS      1 Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirata,M.,
              Suto,Y., Hirai,M., Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
TITLE        Assignment of 118 novel cDNAs of cynomolgus monkey brain to human
              chromosomes
JOURNAL      Gene 275 (1), 31-37 (2001)
MEDLINE      21458551
PUBMED       11574149
REFERENCE    2 (bases 1 to 1998)
AUTHORS      Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
TITLE        Direct Submision
              Submitted (16-OCT-2002) Katsuyuki Hashimoto, National Institute of
              Infectious Diseases, Division of Genetic Resources, 23-1, Toyama
              1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
              (E-mail: khashim@nih.go.jp, URL: http://www.nih.go.jp/yoken/genbank/,
              Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)
COMMENT      Lab host: TOP10
              Vector: PMB18S-FL3 (Acc.No. AB009864)
              R. Site1: DraIII (CACTGTGTG)
              R. Site2: DraIII (CACCAGTGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer
              [ATGTGGCTTTTGTGTGTGT]; double-stranded cDNA was synthesized
              using specific 5' and 3' primers and amplified by PCR. The PCR
              product was digested with SfiI and size selection was performed to
              exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
              into distinct DraIII sites of pMBS-FL3. XhoI sites just outside
              the DraIII sites can be used to isolate the cDNA insert. Libraries
              were constructed by oligo-capping method. Custom primers used for
              sequencing ( 5' and primer [CTTCTCTCTTAAAGCTGCG] ,
              3' end primer [CAGCTGACAGCTCGAGCA] ).
FEATURES
source       1. 1998
              /organism="Macaca fascicularis"
              /mol_type="mRNA"
              /db_xref="taxon:9541"
              /clone="QF1A-17705"
              /sex="male"
              /tissue_type="frontal lobe left"
              /clone_id="macaque brain cDNA library QF1A"
              /dev_stage="adult"
              222..1223
              /codon_start=1
              /product="hypothetical protein"
              /protein_id="BAC21610.1"
              /db_xref="GI:24053729"
              /translation="MWGGGWENGTLVBEIPACQDLQLSLSLGLVGVVGLCY
              NALIVLNIHSHKASMTMPDVYFVNNAVAGLVLSALAPHLGPSSONALMWAGSEVH
              VALQIPFNVSSLVNAYSTALSLDHYIRALPRVTMASVYNTRHVCGFVWGALITSF
              SLILFYICSHVSTRALBCAKMONAATAATFVIGVVPALALYALVLSVRREDT
              PLDRPTGLBPSAHRLIVATYCTORGIMTPHYLVILGHTVILISROPVDAAHYLGILHF
              VMDLSKLAFSSFTPLIRYMNQSPFGKRLRMKKLPCGDHSCSPDHIGQVOLA"
CDS
Query Match 91.4%; Score 915.6; DB 9; Length 1998;
Best Local Similarity 94.6%; Pred. No. 2.7e-148;
Matches 948; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
ORIGIN
      1 ATGTGAGCTGACAGCTGTTCAACGCAAGGCTGTGTGAGAGCTGCTCTGCGAG 60
      222 ATGTGAGGCTGCGGCTGTTCAACGCAAGGCTGTGTGAGAGCTGCTCTGCGAG 281
      61 GACCTGACAGTGGGCTGTCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
      282 GACCTGACAGTGGGCTGTCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 341
      121 GGCCTGTGTGTAACAAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
      342 GGCCTGTGTGTAACAAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 401
      181 ATGCCGAGAGTGTACTTTGTCAACATGGCAGTGGCGGCTGTCTCAGCGCCCTGGCC 240
      402 ATGCCGAGAGTGTACTTTGTCAACATGGCAGTGGCGGCTGTCTCAGCGCCCTGGCC 461

```

```

Oy 241 CCTGTGACCTGTGCGCCCGGAGCTCCCGTGGGGGCTGTGGAGTGTGGCGCGGA 300
Db 462 CCGTGGACCTGTGCGCCCGGAGCTCCCGTGGGGGCTGTGGAGTGTGGCGCGGAG 521
Oy 301 GTCCACGTGGCACTGCAAGATCCCTTCAATGTGCTTCACTGTGTGCATGTATCTCCACC 360
Db 522 GTCCACGTGGCACTGCAAGATCCCTTCAATGTGCTTCACTGTGTGCATGTATCTCCACT 581
Oy 361 GCGCTGTGAGCGCTGACCACTTACATGAGGGTGCATCTGCGCGGACCTTACATGCGCAGC 420
Db 582 GCGCTGTGAGCGCTGACCACTTACATGAGGGTGCATCTGCGCGGACCTTACATGCGCAGC 641
Oy 421 GTGTACAACACGCGGACGTTGTGCGCTTCTGTGGGGTGGCGCGCTGTGACCAAGCTTC 480
Db 642 GTGTACAACACGCGGACGTTGTGCGCTTCTGTGGGGTGGCGCGCTGTGACCAAGCTTC 701
Oy 481 TCCCGGTGCTCTTCACTTATGTGACGCAATGTGTTCACCGCGCGCTTGAAGTGGCCAG 540
Db 702 TCCCGGTGCTCTTCACTTATGTGACGCAATGTGTTCACCGCGCGCTTGAAGTGGCCAG 761
Oy 541 ATGCAAGAACGCAAGAGCTGCGGACGCGGCTGGTGTATTCGCGTACGTTGCGGACGA 600
Db 762 ATGCAAGAACGCAAGAGCTGCGGACGCGGCTGGTGTATTCGCGTACGTTGCGGACGA 821
Oy 601 CTGGCCACCTCTTACGCGCTGTGTGTCTCTCCGCGTCCGACGAGGACACGCGCCCTG 660
Db 822 CTGGCCACCTCTTACGCGCTGTGTGTCTCTCCGCGTCCGACGAGGACACGCGCCCTG 881
Oy 661 GACCGGGACACGCGGCGCGCTGAGCGCTTGGGACACAGGCTGTGGTGGCCACCGTGTGC 720
Db 882 GACCGGGACACGCGGCGCGCTGAGCGCTTGGGACACAGGCTGTGGTGGCCACCGTGTGC 941
Oy 721 AGCGAGTTTGGGCTGTGACGCGCACATATCTGTCTGTGGGGACACGCGTCAATC 780
Db 942 AGCGAGTTTGGGCTGTGACGCGCACATATCTGTCTGTGGGGACACGCGTCAATC 1001
Oy 781 TCGCGAGGGAGAGCGCGTGGACGACACTTCTGGGGCTACTGCACTTGTGTGAAGATTTC 840
Db 1002 TCGCGAGGGAGAGCGCGTGGACGACACTTCTGGGGCTACTGCACTTGTGTGAAGATTTC 1061
Oy 841 TCGCAATCTCTGCTTCTTCCAGACGCTTGTGACACCACTTCTTACCGCTTACATGAAAC 900
Db 1062 TCGCAATCTCTGCTTCTTCCAGACGCTTGTGACACCACTTCTTACCGCTTACATGAAAC 1121
Oy 901 CAGAGCTTCCCGACGACGCTTCCACGCGTGTGAAAGAAAGCTTCTGCGGGACCGCGAC 960
Db 1122 CAGAGCTTCCCGACGACGCTTCCCGCGCGTGTGAAAGAAAGCTTCTGCGGGACCGCGAC 1181
Oy 961 TCGTCCCGGACACATGCGGGTGGACGAGGCTGCGCGTAG 1002
Db 1182 TCGTCCCGGACACATGCGGGTGGACGAGGCTGCGCGTAG 1223

```

```

RESULT 13
AK097629/c 2240 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens CDNA FLJ40310 f18, clone TEST12029672.
DEFINITION AK097629
ACCESSION AK097629
VERSION AK097629.1 GI:21757459
KEYWORDS oligo capping; f18 (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Okamoto, Y., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Ohashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuma, N., Sato, K.,

```

```

Takahashi, M., Yamazaki, M., Nishikawa, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Horita, T., Kusano, O., Kanohori, K., Takahashi, Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiyasu, S., Komai, F., Hara, R.,
Takeshita, K., Arita, M., Imose, N., Muesel, K., Yuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Makebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujizawa, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hiki, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunawa, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hara, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuo, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039
2
JOURNAL
PUBLISHED
REFERENCE
AUTHORS
Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hirooka, S.,
Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Horita, K., Watanabe, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T.,
Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Takahashi, Fujii, A., Oshima, A., Sugiyama, A.,
Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuo, Y., Nagai, K.,
and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2240)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatachi, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Source
1..2240
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST12029672"
/tissue_type="testis"
/clone_lib="TEST12"
/note="cloning vector: PME18SFLU3"
ORIGIN
Query Match 87.6%; Score 877.8; DB 9; Length 2240;
Best local Similarity 99.8%; Pred. No. 9e-142;
Matches 879; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 122 GCGTGTGCTACACGCGCTGTGTGTGCTGCGCACTTACACGACGCGCGATGACCA 181
Db 1977 GCGTGTGCTACACGCGCTGTGTGTGCTGCGCACTTACACGACGCGCGATGACCA 1918
Oy 182 TGCCGACGCTGCTTGTGCAACATGCGACGTGCGACGCTGTGCTTACGCGCGCTGCCCC 241
Db 1917 TGCCGACGCTTACTTGTGCAACATGCGACGTGCGACGCTGTGCTTACGCGCGCTGCCCC 1858

```

```

QY 242 CTGTGACCTGCTGGGCGCCCGGAGCTCCCGGCGGCGCTGCTGAGATGCTGGCGGAGG 301
DB 1857 CCGTACACCTGCTGGGCGCCCGGAGCTCCCGGCGGCGCTGCTGAGATGCTGGCGGAGG 1798
QY 302 TCCACGTGACCTGACAGATCCCTTTCATGCTGCTCACTGCTGAGCCATGATCTCCACCG 361
DB 1797 TCCAGGTGACCTGACAGATCCCTTTCATGCTGCTCACTGCTGAGCCATGATCTCCACCG 1738
QY 362 CCGTGTGAGCTCGACCACTACATGAGCGTGCATCTGCCCGGAGCCTACATGCGCAGCG 421
DB 1737 CCGTGTGAGCTCGACCACTACATGAGCGTGCATCTGCCCGGAGCCTACATGCGCAGCG 1678
QY 422 TGTAAACACGCGGCGACGCTGCGGCTGCTGAGGCTGGCGCGCTGACACAGCTTCT 481
DB 1677 TGTAAACACGCGGCGACGCTGCGGCTGCTGAGGCTGGCGCGCTGACACAGCTTCT 1618
QY 482 CCGTGTGAGCTCGACCACTACATGAGCGTGCATCTGCCCGGAGCCTACATGCGCAGCG 541
DB 1617 CCGTGTGAGCTCGACCACTACATGAGCGTGCATCTGCCCGGAGCCTACATGCGCAGCG 1558
QY 542 TGCAGAACGAGAGAGCTGCGGACCGACGCTGCTTCACTGCGCTACGCTGCTCCAGCAC 601
DB 1557 TGCAGAACGAGAGAGCTGCGGACCGACGCTGCTTCACTGCGCTACGCTGCTCCAGCAC 1498
QY 602 TGGCCACCTCTACACGCGCTGCTGCTCTCCCGGCGGAGGAGGACACGCGCTTGG 661
DB 1497 TGGCCACCTCTCTACACGCGCTGCTGCTCTCCCGGCGGAGGAGGACACGCGCTTGG 1438
QY 662 ACCGAGACACGCGGCGCGCTGAGAGCCCTCGGACACAGAGCTGCTGAGGCGACCGCTGGA 721
DB 1437 ACCGAGACACGCGGCGCGCTGAGAGCCCTCGGACACAGAGCTGCTGAGGCGACCGCTGGA 1378
QY 722 CCGAGTTTGGGCTCTGAGAGCGCCACACTATCTGATCTGCTGAGGAGGACACGCTCATCT 781
DB 1377 CCGAGTTTGGGCTCTGAGAGCGCCACACTATCTGATCTGCTGAGGAGGACACGCTCATCT 1318
QY 782 CCGAGAGGAGAGCCCGTGGAGCGCACTACTGCGGCTGCTGAGGAGGAGGAGGATTTCT 841
DB 1317 CCGAGAGGAGAGCCCGTGGAGCGCACTACTGCGGCTGCTGAGGAGGAGGAGGATTTCT 1258
QY 842 CCAAACTCTGAGCTCTCTGAGAGCTTGTGACACCACTTCTCTACCGCTACATGAGC 901
DB 1257 CCAAACTCTGAGCTCTCTGAGAGCTTGTGACACCACTTCTCTACCGCTACATGAGC 1198
QY 902 AGAGCTTCTCCAGCAAGCTCTCAACGCGCTGATGAAAAGCTGCTGCGGAGCAGCGCAGCT 961
DB 1197 AGAGCTTCTCCAGCAAGCTCTCAACGCGCTGATGAAAAGCTGCTGCGGAGCAGCGCAGCT 1138
QY 962 GCTCCCGGAGCAGCATGCGGCTGCAACAGCTGCTGCGCTGAG 1002
DB 1137 GCTCCCGGAGCAGCATGCGGCTGCAACAGCTGCTGCGCTGAG 1097

```

```

RESULT 14
BC003323 1468 bp mRNA linear ROD 29-JUN-2004
DEFINITION Mus musculus cDNA sequence BC003323, mRNA (cDNA clone MGC:7035
IMAGE:3156060), complete cda.
ACCESSION BC003323
VERSION BC003323.1 GI:13097086
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1468)
Struhsberg, R.L., Peingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schlier, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diachenko, L., Marcus, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Udén, T.B., Toshitaki, S.,

```

```

REMARK
COMMENT
JOURNAL
PUBMED
AUTHORS
TITLE
JOURNAL
1247932
2 (bases 1 to 1468)
Struhsberg, R.
Direct Submission
Submitted (20-FEB-2001) National Institute of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gumartine, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louie, H.,
Kowla, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAX Plate: 6 Row: m Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13385047.
Location/Qualifiers
1. 1468
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/db_xref="MGC:7035 IMAGE:3156060"
/tissue_type="mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_id="NCI CGAP_Mam6"
/lab_host="DH10B"
/notes="Vector: pCMV-Sport6"
1. 1468
/gene="BC003323"
/notes="aynomy: MGC7035"
/db_xref="LocusID:80290"
/db_xref="MGI:1933113"
72..1073
/gene="BC003323"
/codon_start=1
/product="cDNA sequence BC003323"
/protein_id="AAH03323.1"
/db_xref="GI:13097087"
/db_xref="LocusID:80290"
/db_xref="MGI:1933113"
/translation="MMSGCPPLNSTMAEPRICRNIRIGIMTSLIYLNGVPSIGYN
ALLVLANLAKNTMTMPDYFVNMAVAGIVLPAVLIGFHSRWLMSSEAHV
TLLILFNVAISLSTALSLIDYIERALPRTYASVYNTBHVCGFPWGAIVLIFS
SLIFVIGSVSRISAECAKMTBADAIIIVIGVVGAVLAVLALISRIKEDPT
LDQDSRLDPSVHRLVAVTCQFGLMPYVSLGHTVYLSGRVTEGHYGLIIVAK
DLAKFLAPSSSVPLVLYINKAPGKLRMLKMKHGHGRRHCSDPDSIQVMAQA"

```

ORIGIN

```

Query Match      60.9%; Score 610; DB 10; Length 1468;
Best Local Similarity 76.9%; Pred. No. 1.5e-95;
Matches 771; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

OY 1 ATGTGGAGCTGACGTGTTCAACAGGACAGGAGCTGTGAGAGAGCTGCTCTCCAGC 60
DB 72 ATGTGGAGCTGTGGCCCACTCAACAGACAGAGCTGGAGAGAGCCG---TGTCGGC 128
OY 61 GACCTGACAGCTGGGCTGTCACTGTGTGCTGCTGAGGCTGTGAGTGGGCTGCCAGT 120
DB 129 AACCTGCGCTGGGGCTGTGGGTTCCTGCTGCTCTTACTCTGGGGGACAGGGTCCCTGG 188
OY 121 GACCTGTGTACAGCCCTGTGGGTGCTGACCACTGACAGCAAGGACAGCAATGCC 180
DB 189 AGCTTAAGCTTAATAGCTTTCTTGAGCTGGGCAACTGGCCAGCAAGAAACACATGACC 248
OY 181 ATGCGGAGCTGTACTTTGTCAACATGAGCAGTGGAGGAGGCTGTGCTCAGCCCTGGCC 240
DB 249 ATGCCGAGCTGTACTGTGTGAACATGGCCGTGGGGGCTGTGCTCAGGCACTGGCA 308
OY 241 CTTGTGACACTCTGTGGCCCCCGAGCTCCGGTGGGCGCTGTGAGTGTGGGCGGAA 300
DB 309 CTTGTGACTGTGTGGGCTGCTGCCCACTCAAGTGGGCTGTGAGGCTCAGCAGTGA 368
OY 301 GTCCAGTGTGACATGATCCCTTCAATGATGTCCTCACTGGAGGCAATGATCTCCAC 360
DB 369 GCCCATGTGACACTCTCACTCTGTTCATCAAGTGGCTTCCCTGTGTACCAATGATCTCACT 428
OY 361 GCGCTGTGAGCTGTGACCACTACATCGAGCTGTGACTGCGGACCTACATGAGCCAGC 420
DB 429 GCACTGTGAGCTGTGACTACATCGAGCGGTGCTGTGCAAGCACTACATGAGCCAGT 488
OY 421 GTGTCAACACCGGCGACAGTGTGCGGCTTCTGTGGGTGGCGGCTGTGACCACTTC 480
DB 489 GTGTCAACACCGGCGACAGTGTGCGGCTTCTGTGGGTGGCGGCTGTGACCACTTC 548
OY 481 TCTGTGCTGTCTTTTACATCTGACAGCAGTGTGTCACCGCGGCTGTGAGTGGGCGCAG 540
DB 549 TCTGTGCTGTCTTTTACATCTGACAGCAGTGTGTCACCGCGGCTGTGAGTGGGCGCAG 608
OY 541 ATGCAAGACGAGAGCTGCGGACCGCAAGCTGTGTTCATCGGCTACGAGTGGCCAGCA 600
DB 609 ATGCAAGACGAGAGCTGCGGACCGCAAGCTGTGTTCATCGGCTACGAGTGGCCAGT 668
OY 601 CTGGGACCTCTTACAGCGCTGTGTACTTTCGCCGCTGCCAGAGGAGACAGCCCTTG 660
DB 669 CTGGGCTGTGTATGCTTGTGCACTCACTTTCGAGAAATCGGGAAGAGACACACCCCTG 728
OY 661 GACGCGGACAGGCGCTGTGAGGCTTGTGGCAACAGAGCTGTGTGGCCAGCGTGTGC 720
DB 729 GACGCGGACAGGAGGCTGTGAGGCTTGTGGCAACAGAGCTGTGTGGCCAGCGTGTGC 788
OY 721 AGGCAATTTGGAGCTGTGAGCGCACACTATGTGATCTGTGAGGAGCACAGGTATATC 780
DB 789 ACTAGTTTGGCTCTGTGACACACTTACTTACTTGAAGC---CTGGGCGACACAGTGTGACG 845
OY 781 TCGGAGGAGGAGCCGCTGTGAGGACGACACTTGTGGGCTGTGACACTTGTGTAAGATTC 840
DB 846 TCGAGGAGGAGGAGCCGTGTGAGGAGGAGGATTTATGTGGCACTTCAAGTGTCTTAAGAGCTTG 905
OY 841 TCCAAACTCTGTGAGCTTGTGAGAGCTTGTGACACCACTTCTTACCGCTACATATGAC 900
DB 906 GCCAAGTTCCTGTGGCTTCTCAAGCAAGTTCGTGTGACACCACTGCTCTTACCTTACATCAAC 965
OY 901 CAGAGCTTCCGACAGCTTCAACAGGCTGTGATGAAAAAGCTGCGCTGGGAGCGGAGC 960
DB 966 AAACCTTTCCTCCGCAAGCTCCGGCGGCTGTATGAAAGATGACCTGGGGCGCGGCAC 1025
OY 961 TGTCTCCCGGACCAATGAGGGGTGTGACAGAGTGTGTGGCTGTG 1002
DB 1026 TGTCTCCCGGACCTCTCGGGGATACAGCAGGTGTATGGACAG 1067

```

```

RESULT 15
AC140216
LOCUS
DEFINITION
AC140216 194560 bp DNA linear ROD 10-JUL-2004
Mus musculus BAC clone RP23-276B20 from chromosome 5, complete
sequence.
AC140216
AC140216.3 GI:48675535
HTG.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.
1 (bases 1 to 194560)
Berghoff, A., Cotton, M., Bielicki, L. and Haglund, K.
The sequence of Mus musculus BAC clone RP23-276B20
2 (bases 1 to 194560)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 194560)
Wilson, R.K.
Direct Submission
Submitted (08-MAY-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 194560)
Wilson, R.K.
Direct Submission
Submitted (12-JUN-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 194560)
Wilson, R.K.
Direct Submission
Submitted (10-JUL-2004) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jun 12, 2004 this sequence version replaced gi:47084696.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Summary Statistics
Center project name: M_BA0276B20
-----

```

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:

The RPCI-23 BAC library has been constructed by Kazuhiro Osegawa
 and Minko Tareno in the laboratory of Pieter de Jong
 (http://www.chori.org) from female C57BL/6J mouse kidney and/or
 brain genomic DNA. The clone and detailed information can be
 obtained from Research Genetics, Inc. (http://www.resgen.com) or
 Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC144902.

FEATURES

source
Location/Qualifiers

1..194560
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP23-276B20"
/clone_11b="RPCI-23"
388..741
repeat_region
/rpt_family="MaLR"
742..1371
repeat_region
/rpt_family="MaLR"
1513..1692
repeat_region
/rpt_family="B2"
1736..1795
repeat_region
/rpt_family="MaLR"
1861..2178
repeat_region
/rpt_family="MaLR"
2181..2410
repeat_region
/rpt_family="MaLR"
4755..5099
unsure
/note="Unresolved simple sequence repeat."
repeat_region
6703..6811
/rpt_family="Alu"
7493..7637
repeat_region
/rpt_family="B4"
7724..7845
repeat_region
/rpt_family="B4"
7973..8026
repeat_region
/rpt_family="ZP3AR_MM"
8027..8108
repeat_region
/rpt_family="ID"
8109..8130
repeat_region
/rpt_family="ZP3AR_MM"
8131..8213
repeat_region
/rpt_family="ID"
8214..8272
repeat_region
/rpt_family="ZP3AR_MM"
8312..8591
repeat_region
/rpt_family="B4"
8679..9004
repeat_region
/rpt_family="Alu"
9367..9461
repeat_region
/rpt_family="Alu"
9469..9605
repeat_region
/rpt_family="Alu"
9926..9951
repeat_region
/rpt_family="ERV1"
9952..10045
repeat_region
/rpt_family="Alu"
10046..10107
repeat_region
/rpt_family="ERV1"
10109..10251
repeat_region
/rpt_family="Alu"
10254..10387
repeat_region
/rpt_family="Alu"
10418..10520
repeat_region
/rpt_family="B4"
10521..10650
repeat_region
/rpt_family="Alu"
10701..10857
repeat_region
/rpt_family="B4"
11131..11188
repeat_region
/rpt_family="ZP3AR_MM"
11189..11272
repeat_region
/rpt_family="ID"
11273..11294
repeat_region
/rpt_family="ZP3AR_MM"
11295..11377
repeat_region

repeat_region
/rpt_family="ID"
11378..11399
repeat_region
/rpt_family="ZP3AR_MM"
11400..11482
repeat_region
/rpt_family="ID"
11483..11504
repeat_region
/rpt_family="ZP3AR_MM"
11505..11587
repeat_region
/rpt_family="ID"
11588..11639
repeat_region
/rpt_family="ZP3AR_MM"
12030..12127
repeat_region
/rpt_family="L1"
12866..12934
repeat_region
/rpt_family="ID"
12946..13155
repeat_region
/rpt_family="MER1_type"
13156..13229
repeat_region
/rpt_family="ZP3AR_MM"
13231..13309
repeat_region
/rpt_family="ID"
13326..13448
repeat_region
/rpt_family="B4"
13514..13721
repeat_region
/rpt_family="B2"
13747..13903
repeat_region
/rpt_family="B4"
13904..14032
repeat_region
/rpt_family="L1"
14089..14196
repeat_region
/rpt_family="B4"
14372..14515
repeat_region
/rpt_family="B4"
14920..15033
repeat_region
/rpt_family="MaLR"
15274..15420
repeat_region
/rpt_family="Alu"
15610..15664
repeat_region
/rpt_family="B2"
15696..15746
repeat_region
/rpt_family="Alu"
15748..15914
repeat_region
/rpt_family="L1"
15915..16104
repeat_region
/rpt_family="B2"
16105..16252
repeat_region
/rpt_family="L1"
16261..16535
repeat_region
/rpt_family="B4"
16609..16737
repeat_region
/rpt_family="L1"
16983..17245
repeat_region
/rpt_family="B4"
17302..17447
repeat_region
/rpt_family="Alu"
17475..17603
repeat_region
/rpt_family="Alu"
17992..18051
repeat_region

Query Match 60.9%; Score 610; DB 10; Length 194560;
Best Local Similarity 76.9%; Pred. No. 8.6e-96;
Matches 771; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

QY 1 ATGTGAGCTGCAGCTGTTCAACGACACAGGCTGTGAGAGAGCTGCTGCTGCAG 60
DB 97303 ATGTGAGCTGTGCCCCACTCAACAGACAGCGTGGCTGAGAGAGCCG---TGTGCCGG 97359
QY 61 GACCTGCAGCTGGGCTGTCACTGTGTGCTGCTGGGCTGTGTGGAGCTGCCAGT 120
DB 97360 AACCTGGCGCTGGGCTGTGGCTCTTCGCTGCTTACCTGGGAGCAGGAGCTCCCTGG 97419
QY 121 GGCTGTGCTACAGCGCTGTGCTGTGCTGCAACTACACAGCAAGGCGCATGACC 180

Db 97420 AGCTTAGCTATATGCTCTTCTGCTGCGCAACTGAGCAAGACACATGACC 97479
QY 181 ATGCCGAGAGTGTACTTTGTCAACATGAGCAAGGAGGCGCTGTGCTCAGCGCCCTGACC 240
Db 97480 ATGCCGAGAGTGTACTTTGTCAACATGAGCAAGGAGGCGCTGTGCTCAGCGCACTGACA 97539
QY 241 CCTGTGCACTGTGCGGCCCCCGAGCTCCCGGTGGCGCTGTGAGGTGTGGCGGAGAA 300
Db 97540 CCGTGGTACCTGTTGGGCCCTGGCCACTCAAGGTGGGCCCTGTGGAGCCTCAGCAGTGAG 97599
QY 301 GTCCACGTGGCACTGCAAGATCCCTTCANGTGTCTCACTGTGGCCATGTACTCCACC 360
Db 97600 GCGCATGTGACACTGTCTATCTCTGTCAAGGTGGCTTCCCTGTGACCATGTACTCCACT 97659
QY 361 GCCCTGTGAGCCTTGACCACTAATCAGAGGTGCACTGCGCGGACCTTACATGAGCCAGC 420
Db 97660 GCACGTGCTGAGCCTTGACTACTATCATGAGGTGCTCCCTGCCACGACACTTACATGAGCCAGT 97719
QY 421 GTGTACAAACAGCGCGGCAAGTGTGCGGCTTGTGTGGGGCGCGCTGTGACCAAGCTTC 480
Db 97720 GTGTACAAACAGCGCGGCAAGTGTGCGGCTTGTGTGGGGCGCGGCTGTGACCAAGCTTC 97779
QY 481 TCCTGTGCTGTCTTCTACATCTGTGAGCCATGTGCCACCGCGGCTAAGGTGCGCAAG 540
Db 97780 TCCTGTGCTGTCTTCTACATCTGTGAGCCATGTGCCACCGCGGCTAAGGTGCGCGG 97839
QY 541 ATGAGAAACGAGAAAGCTGCCAGCGCAAGCTGTGTTCATCCGCTACGAGGTGCCAGCA 600
Db 97840 ATGAGAAACGAGAGGAGCGAGCGCAAGCTGTGTTCATCCGCTACGAGGTGCCAGGT 97899
QY 601 CTGGCCACCTCTACGCGGCTGTGTACTCTCCCGGCTCCGAGGAGAGCAAGCCCTTG 660
Db 97900 CTGGCTGTGTGTATGCTCTGAGCACTCATCTCGAAGATCGGAGAGAGACACCCCTG 97959
QY 661 GACCGGAGCAAGGCGCGGCTGAGGCCCTGAGCAACAGGCTGTGTGGCCACCGGTGTC 720
Db 97960 GACCAAGAGCACAGAGGCTGAGGCCCTCGGTGCAAGGCTGTGTGGCCACCGGTGTC 98019
QY 721 ACCGAGTTGGGCTTGGAAGCCCACTATCTGATCTGTGTGGGCAACGGTCTATCATC 780
Db 98020 ACTCATTTGGCTCTGGAACCTTACTTGTGAGC--CTGGGGCACACAGTGTGACG 98076
QY 781 TCGCGAGGAGAGCCGCTGAGCGCACACTACTGTGGGCTACTGCACTTTGTGAAGATTTC 840
Db 98077 TCAAGGGGAGAGCCGCTGAGGGGCACTTATCTGGGCACTCTACAGTTGCTAAGGACCTG 98136
QY 841 TCGAAGCTCTGGCTTCTCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
Db 98137 GCTAAGTCTCTGACCTTCTCAAGCAGTTCTGTGACACCACTGCTTACCGTTACATCAAC 98196
QY 901 CAGAGCTTCCCGAGAGGCTCCAGCGGCTGATGAAAAGCTGCCCTGTGGGGAGCCGGGAC 960
Db 98197 AAAAGCTTCCCGGAGAGGCTCCGGGGGCTGATGAAAAGATGCACTGCGGGGCGCGGCAC 98256
QY 961 TGCTCCCCGAGACCATGTGGGGGTGAGGAGGTGCTGGCGTAG 1002
Db 98257 TGCTCCCCGAGACCCCTCGGGGATACAGAGGATGATGGACAG 98298

Search completed: December 15, 2004, 00:42:49
Job time : 3379.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:00:04 ; Search time 469.273 Seconds
(without alignments)
11208.653 Million cell updates/sec

Title: US-09-995-225b-1

Perfect score: 1002
Sequence: 1 eegtgagagctgcgcgtcgtc.....tcgacgsgtgcgtgcgcgtag 1002

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23sep04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1002	100.0	1002	6	ABT04866 Human G P
2	1002	100.0	1002	6	ABK6286 Human TGR
3	1002	100.0	1402	10	ADC86562 Human G P
4	1002	100.0	1778	4	AA642858 Human G P
5	1002	100.0	1785	8	AB242598 Human G P
6	1002	100.0	1785	12	ADC30022 Human G P
7	1002	100.0	1926	8	ACCT0841 Human G-P
8	1002	100.0	1928	8	ACA61176 Human G-P
9	1000.4	99.9	1614	3	AA290175 Human che
10	1000.4	99.8	1750	4	AA642836 Human G P
11	1000.4	99.8	1878	6	ABX73053 Human G-P
12	1000.4	99.8	1893	4	AAH99600 Human pro
13	999	99.7	2216	8	ACCT0860 Insert CD
14	999	99.7	2258	8	ACCT0859 Insert CD
15	999	99.7	2363	8	ACCT0861 Insert CD
16	989.4	98.7	1879	3	AA290523 Human GPC
17	610	60.9	1468	12	AD030312 Mouse GPC
18	577	57.6	2651	4	ABK42895 Genomic s
19	577	57.6	2651	9	ADB61051 Connectiv
20	521	52.0	1269	6	ABA04303 Human ops
21	422.2	42.1	559	6	ABN26732 Human ORF

22	358.6	35.8	398	9	ACH20905
23	350.6	35.0	461	9	ACH43678
24	343.8	34.3	607	4	AA182081 Human pol
25	304	30.3	497	6	ABX73103 Mouse var
26	275	27.4	275	6	ABX73076 Human GPC
27	236.4	23.6	238	6	ABX73074 Human GPC
28	229.4	22.9	300	3	AA290174 Genomic s
29	178	17.8	930	4	ABK42894 Fragment
30	178	17.8	930	6	ADB61050 Connectiv
31	169	16.9	294	6	ABX73105 Rat varia
32	166	16.6	259	6	ABX73075 Human GPC
33	162.6	16.2	266	6	ABX73104 Rat varia
34	134	13.4	255	6	ABX73071 Human GPC
35	132	12.2	281	6	ABX73072 Human GPC
36	89.4	8.9	363	6	ABX73073 Human GPC
37	71	7.1	2932	6	AA226878 Human G-P
38	69.8	7.0	1089	3	AAA30581 Human G P
39	69.8	7.0	1089	6	AA226832 Human G-P
40	69.8	7.0	1089	8	ACCT8113 Human G P
41	69.8	7.0	1089	10	ADC22522 Human G P
42	69.8	7.0	1089	10	ADH13995 Human GPR
43	69.8	7.0	1089	12	AD029766 Human gen
44	69.8	7.0	1365	6	AB234905 Human nov
45	69.8	7.0	1365	10	AD684894 Farnesyl

ALIGNMENTS

RESULT 1	ABT04866	ABT04866 standard; cDNA; 1002 BP.
ID	ABT04866	
AC	ABT04866	
XX		
DT	11-OCT-2002 (first entry)	
DE	Human G protein coupled receptor hrup28 coding sequence.	
KW	Human; G-protein coupled receptor; GPCR; hrup28; hrup29; hrup30; hrup31; hrup32; hrup33; hrup34; hrup35; hrup36; hrup37; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200242461-A2.	
XX		
PD	30-MAY-2002.	
XX		
PF	26-NOV-2001; 2001WC-US044386.	
XX		
PR	27-NOV-2000; 2000US-0253404P.	
PR	12-DEC-2000; 2000US-0253466P.	
PR	20-FEB-2001; 2001US-0270266P.	
PR	20-FEB-2001; 2001US-0270286P.	
PR	06-APR-2001; 2001US-0282032P.	
PR	06-APR-2001; 2001US-0282356P.	
PR	06-APR-2001; 2001US-0282358P.	
PR	06-APR-2001; 2001US-0282365P.	
PR	14-MAY-2001; 2001US-0290917P.	
PR	31-JUL-2001; 2001US-0309208P.	
XX		
PA	(AREN-) ARENA PHARM INC.	
XX		
PI	Chen R, Chu ZL, Deng HT, Lowitz KP, Pride C;	
XX		
DR	WPI; 2002-566565/60.	
DR	P-PSDB; ABJ04068.	
XX		
PT	Novel endogenous and non-endogenous versions of G protein-coupled receptor useful for identification of candidate compounds as receptor agonists or antagonists for use as therapeutic agents.	
XX		
PS	Claim 3; Page 54-55; 84pp; English.	

XX The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a GPCR coding sequence of the invention

Sequence 1002 BP; 163 A; 333 C; 313 G; 193 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1002;	DB 6;	Length 1002;
Best Local Similarity	100.0%;	Pred. No. 3.5e-194;		

Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGTGAGCTGAGCTGTTTCAACGGCACAAGGCTGTGTGAGAGAGCTGCTGCTCCAG	60
Dp	1	ATGTGAGCTGAGCTGTTTCAACGGCACAAGGCTGTGTGAGAGAGCTGCTGCTCCAG	60
QY	61	GACCTGCAGCTGGGGCTGTCACTGTTTGCCTGTGGGCTGGTGGGCGTGCAGTG	120
Dp	61	GACCTGCAGCTGGGGCTGTCACTGTTTGCCTGTGGGCTGGTGGGCGTGCAGTG	120
QY	121	GGCCGTGTCTAACAAGCCCTGTGTGTGTGGCCAACTTACAAGCAAGCCAGCATGACC	180
Dp	121	GGCCGTGTCTAACAAGCCCTGTGTGTGTGGCCAACTTACAAGCAAGCCAGCATGACC	180
QY	181	ATGCCGAGAGTGTACTTTGTCAACATGCGAGTGGACAGGCGCTGGTGTCAAGCGCCCTGGCC	240
Dp	181	ATGCCGAGAGTGTACTTTGTCAACATGCGAGTGGACAGGCGCTGGTGTCAAGCGCCCTGGCC	240
QY	241	CCTGTGCACCTGCTGGGCCCCCGAGCTCCCGGTGGCGCTGTGAGTGTGGCGGCGAA	300
Dp	241	CCTGTGCACCTGCTGGGCCCCCGAGCTCCCGGTGGCGCTGTGAGTGTGGCGGCGAA	300
QY	301	GTCCACGTGGGCACTCAGATCCCTTCAATGTGTCTCATGTGGGCCAATGATCTCAC	360
Dp	301	GTCCACGTGGGCACTCAGATCCCTTCAATGTGTCTCATGTGGGCCAATGATCTCAC	360
QY	361	GCCCTGTGAGCTCTGACCACTACATGAGCGGTGACCTGCGCGACCTTACATGGCAGAC	420
Dp	361	GCCCTGTGAGCTCTGACCACTACATGAGCGGTGACCTGCGCGACCTTACATGGCAGAC	420
QY	421	GTGTACAACAGCGCGGCACTGTGTGCGGCTTGTGTGTGGGTGGGCGCTGTGACACACTTC	480
Dp	421	GTGTACAACAGCGCGGCACTGTGTGCGGCTTGTGTGTGGGTGGGCGCTGTGACACACTTC	480
QY	481	TCTCTGCTCTCTTCTACATCTGCAGCCATGTGTCAACCCTGGCGCTAGAATGGCGCAAG	540
Dp	481	TCTCTGCTCTCTTCTACATCTGCAGCCATGTGTCAACCCTGGCGCTAGAATGGCGCAAG	540
QY	541	ATGCAGAAACGAGAAAGTGTCCGACGCGCAGCTGTTCATGTGCTACGTGTGCACGA	600
Dp	541	ATGCAGAAACGAGAAAGTGTCCGACGCGCAGCTGTTCATGTGCTACGTGTGCACGA	600
QY	601	CTGGCGAACCTCTACGCGCTGGTGTATCTCCCGGCGTCCGAGAGGAGACAGCGCCCTG	660
Dp	601	CTGGCGAACCTCTACGCGCTGGTGTATCTCCCGGCGTCCGAGAGGAGACAGCGCCCTG	660
QY	661	GACCGGAGACAGGCGCGGCTGGAAGCCTCGGACACAGAGCTGTGTGGCCACGCTGTGC	720
Dp	661	GACCGGAGACAGGCGCGGCTGGAAGCCTCGGACACAGAGCTGTGTGGCCACGCTGTGC	720
QY	721	ACGAGATTTGGGCTCTGTGACGCGCACATATCTGATCTGTGTGGGCAACAGGTCATCATC	780
Dp	721	ACGAGATTTGGGCTCTGTGACGCGCACATATCTGATCTGTGTGGGCAACAGGTCATCATC	780
QY	781	TCGCGAGGGAAGCCCGTGGAGCGACACTACCTGTGGGCTACCTGTAAGGATTTTC	840
Dp	781	TCGCGAGGGAAGCCCGTGGAGCGACACTACCTGTGGGCTACCTGTAAGGATTTTC	840
QY	841	TCCAAATCTCTGGGCTTCTCCAGACGCTTTGTGACCACTTCTCTACCGCTCACATGAAC	900
Dp	841	TCCAAATCTCTGGGCTTCTCCAGACGCTTTGTGACCACTTCTCTACCGCTCACATGAAC	900

Qy 90 CAAGAGTTCCCGACAGCTCCACGAGGCTAGTAAAAAGTGCCTCGCGGGACCGGAC 960

Db 901 CAAAGTTCCCGACAGCTCCACGAGCTGATGAAAAAGTGCCTCGCGGGACCGGAC 960

Qy 961 TGTCTCCCGGACCAATGGGGGTGACAGCAGTGTCTGCGGTAG 1002

Db 961 TGTCTCCCGGACCAATGGGGGTGACAGCAGTGTCTGCGGTAG 1002

RESULT 2
ABK86286
ID ABK86286 standard; cDNA; 1002 BP.

AC ABK86286;

DT 27-AUG-2002 (first entry)

Human TGR339 cDNA.

KW Human; TGR339; gene; ss; G-protein coupled receptor; GPCR; TGR;

KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;

sleep disturbance; temperature regulation; blood pressure; hypothalamus; KW

3 X

XXI

FT	CDS	1. .1002
FE		744

FT [/product= "Human TGR339 protein"](#)

PN W0200242458-A2.

PD 30-MAY-2002

PF 21-NOV-2001; 2001WO-US043404.

PR 22-NOV-2000; 2000US-0252841P.
PR 22-NOV-2000; 2000US-0252841P.

PR 12-JAN-2001; 2001US-0261377P.
DB 28 JAN 2001 00270EAD

PR 29-MAR-2001; 2001US-0280696P.
 YY

PA (TULA-) TULARIK INC.
VY

PI Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupte JS;

DR WPI; 2002-463633/49.

XX
BT
New York, N.Y.
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2618
2619

PT diagnosis and treatment of diseases such as renal failure, nephritis, hematuria, diabetes and disturbances of electrolyte and acid-base balance.

XX
PS
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840

CC XX Fb00 11/11/2011 11:11:11 (CCCC)

termed TGR, and its associated nucleic acid. The sequences of the invention are useful for identifying a compound that modulates signal

transduction and for identifying a mammal having a TGR-associated disorder. The proteins and nucleic acids are useful in diagnosing and

CC treatment of diseases or conditions such as renal failure, nephritis, hypertension, hyponatremia, retinitis pigmentosa, growth disorders

diabetes insipidus, hyperprolactinaemia and disturbances of thirst, sleep temperature regulation, appetite, blood pressure or any other

syndrome or disease associated with the hypothalamus. The sequences can be used in regulation of circadian rhythms for use as genetic markers

CC for the identification of mutations associated with diseases resulting
CC from GPCR inactivation in particular cell types and for identification of

modulators of GPCR signal transduction. This sequence represents cDNA encoding the human TGR339 polypeptide

xx Sequence 1002 BP; 163 A; 333 C; 313 G; 193 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 1002; DB 6; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 3.5e-194;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGTGAGCTGACGCTGTTTCAACGACAGGGCTGTGTGAGAGAGCTGCTGCTGCCAG 60
DB 1 ATGTGAGCTGACGCTGTTTCAACGACAGGGCTGTGTGAGAGAGCTGCTGCTGCCAG 60
QY 61 GACCTGACAGCTGGGGCTGTCACTGTTGTGCTGTGAGGCTGTGTGGCGCTGCCAGT 120
DB 61 GACCTGACAGCTGGGGCTGTCACTGTTGTGCTGTGAGGCTGTGTGGCGCTGCCAGT 120
QY 121 GGCCTGTGTCAACAGCCCTGCTGTGTGTGCTGCGCAACTTACACAGCAGGCCATGACC 180
DB 121 GGCCTGTGTCAACAGCCCTGCTGTGTGTGCTGCGCAACTTACACAGCAGGCCATGACC 180
QY 181 ATGCCGGAAGTGTACTTTGTCAACATGGAGTGGCAGGCTGTGTGCTGACCGCCCTGGGC 240
DB 181 ATGCCGGAAGTGTACTTTGTCAACATGGAGTGGCAGGCTGTGTGCTGACCGCCCTGGGC 240
QY 241 CCTGTGACCTGTGCGGCCCGCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGCGAA 300
DB 241 CCTGTGACCTGTGCGGCCCGCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGCGAA 300
QY 301 GTCCAGTGTGACCTGTGACATGATCCCTTCAATGTGTCTCACTGGTGGCCATGATCTCAC 360
DB 301 GTCCAGTGTGACCTGTGACATGATCCCTTCAATGTGTCTCACTGGTGGCCATGATCTCAC 360
QY 361 GGCCTGTGAGCTGTGACCACTACATGAGCGGAGCTGCGCGCACTACATGAGCGCAC 420
DB 361 GGCCTGTGAGCTGTGACCACTACATGAGCGGAGCTGCGCGCACTACATGAGCGCAC 420
QY 421 GTGTCAACACCGGACAGCTGTGTGCGGCTTGTGTGGGGTGGGCGCTGTGACCACTTC 480
DB 421 GTGTCAACACCGGACAGCTGTGTGCGGCTTGTGTGGGGTGGGCGCTGTGACCACTTC 480
QY 481 TCTCGCTGCTTTTCTACATCTGACAGCCATGTGTCCACCGCGCTAGAGTGGCGCAG 540
DB 481 TCTCGCTGCTTTTCTACATCTGACAGCCATGTGTCCACCGCGCTAGAGTGGCGCAG 540
QY 541 ATGCAAGACGCAAGCTGTGCGGACCGCAAGCTGTGTATGAGGCTACCTGTGTGCCAGA 600
DB 541 ATGCAAGACGCAAGCTGTGCGGACCGCAAGCTGTGTATGAGGCTACCTGTGTGCCAGA 600
QY 601 CTGGCACCCTCTACGCGCTGTGTGCTACTCTCCCGGCTCCGACAGGAGACAGCCCTTG 660
DB 601 CTGGCACCCTCTACGCGCTGTGTGCTACTCTCCCGGCTCCGACAGGAGACAGCCCTTG 660
QY 661 GACCGGAGACGCGGCGGCTGTGAGCCCTTGGGACACAGGCTGTGTGGCCACCGTGTGC 720
DB 661 GACCGGAGACGCGGCGGCTGTGAGCCCTTGGGACACAGGCTGTGTGGCCACCGTGTGC 720
QY 721 AGCGAGTTGGGCTGTGAGCGCCACACTATCTGCTGTGGGGACACAGCTCATCATC 780
DB 721 AGCGAGTTGGGCTGTGAGCGCCACACTATCTGCTGTGGGGACACAGCTCATCATC 780
QY 781 TCGCGAGGGAAGCCCGTGTGACGACACTACTGTGGGCTACTGCACTTTGTGAAGATTTC 840
DB 781 TCGCGAGGGAAGCCCGTGTGACGACACTACTGTGGGCTACTGCACTTTGTGAAGATTTC 840
QY 841 TCCAAATCTCTGGCTTCTTCTCAGCAGCTTTGTGACACCACTTCTTACCGCTCATGAAC 900
DB 841 TCCAAATCTCTGGCTTCTTCTCAGCAGCTTTGTGACACCACTTCTTACCGCTCATGAAC 900
QY 901 CAGAGCTTCCCGACGACAGCTCCAAAGGCTGTGAAGAAAACCTGCTCGCGGACCGGCAC 960
DB 901 CAGAGCTTCCCGACGACAGCTCCAAAGGCTGTGAAGAAAACCTGCTCGCGGACCGGCAC 960
QY 961 TCTCTCCCGGACCACTGTGGGGTGTGACAGTGTCTGGCTAG 1002
DB 961 TCTCTCCCGGACCACTGTGGGGTGTGACAGTGTCTGGCTAG 1002

```

DB 961 TCTCTCCCGGACCACTGTGGGGTGTGACAGTGTCTGGCTAG 1002

RESULT 3
 AD86562
 ID AD86562 standard; DNA; 1402 BP.

AD86562;
 01-JAN-2004 (first entry)

Human GPCR gene SEQ ID NO:1015.

de; gene; human; GPCR;
 guanosine triphosphate-binding protein coupled receptor; gene therapy.

Homo sapiens.

EP1270724-A2.

02-JAN-2003.

18-JUN-2002; 2002EP-00013517.

18-JUN-2001; 2001JP-00246789.

(NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

(ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATION.

Suwa M, Arai K, Akiyama Y, Aburatani H;

WPI: 2003-315783/31.

P-PDB; AD86563.

New polynucleotide, useful for preparing a composition for treating a

patient in need of increased or suppressed activity or expression of the

guanosine triphosphate-binding protein coupled receptor.

Claim 1: SEQ ID NO 1015; 28pp; English.

The invention relates to a novel polynucleotide encoding a guanosine

triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of

the invention may have a use in gene therapy. The polynucleotide and

polypeptide are useful for preparing a composition for treating a patient

in need of increased or suppressed activity or expression of the

CC guanosine triphosphate-binding protein coupled receptor. The

CC polynucleotide sequences shown in AD85548-AD87616 encode GPCR's of the

invention.

Sequence 1402 BP; 238 A; 456 C; 428 G; 280 T; 0 U; 0 Other;

Query Match 100.0%; Score 1002; DB 10; Length 1402;

Best Local Similarity 100.0%; Pred. No. 3.7e-194;

Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGTGAGCTGACGCTGTTTCAACGACAGGGCTGTGTGAGAGAGCTGCTGCTGCCAG 60
DB 201 ATGTGAGCTGACGCTGTTTCAACGACAGGGCTGTGTGAGAGAGCTGCTGCTGCCAG 260
QY 61 GACCTGACAGCTGGGGCTGTCACTGTTGTGCTGTGAGGCTGTGTGGCGCTGCCAGT 120
DB 261 GACCTGACAGCTGGGGCTGTCACTGTTGTGCTGTGAGGCTGTGTGGCGCTGCCAGT 120
QY 121 GGCCTGTGTCAACAGCCCTGCTGTGTGTGCTGCGCAACTTACACAGCAGGCCATGACC 180
DB 321 GGCCTGTGTCAACAGCCCTGCTGTGTGTGCTGCGCAACTTACACAGCAGGCCATGACC 180
QY 181 ATGCCGGAAGTGTACTTTGTCAACATGGAGTGGCAGGCTGTGTGCTGACCGCCCTGGGC 240
DB 381 ATGCCGGAAGTGTACTTTGTCAACATGGAGTGGCAGGCTGTGTGCTGACCGCCCTGGGC 240
QY 241 CCTGTGACCTGTGCGGCCCGCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGCGAA 300
DB 241 CCTGTGACCTGTGCGGCCCGCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGCGAA 300

```

```

Db      441  CCTGAGACCTGTCGGGCCCCGAGCTCCGGTGGGCGCTGTGAGTGTGGGCGGAA 500
Qy      301  GTCCACGTGGGACCTGCGAGATCCCTTGAATGTCTCACTGTGGCCATGTATCCACC 360
Db      501  GTCCACGTGGGACCTGCGAGATCCCTTGAATGTCTCACTGTGGGCGCATGTACTCACC 560
Qy      361  GCCCTGTGAGCTCGACACTACATGAGGTCATGCGCGGACACTACATGAGCCAGC 420
Db      561  GCCCTGTGAGCTCGACACTACATGAGGTCATGCGCGGACACTACATGAGCCAGC 620
Qy      421  GTGTACAACAACGCGGACAGTGTGGGCTGTGTGGGTCGCGCTGACCAAGCTTC 480
Db      621  GTGTACAACAACGCGGACAGTGTGGGCTGTGTGGGTCGCGCTGACCAAGCTTC 680
Qy      481  TCCCTGGTGTCTTCTACATCTGTGAGGACATGTCTCCACCGGCGCTAGATGCGCCAG 540
Db      681  TCCCTGGTGTCTTCTACATCTGTGAGGACATGTCTCCACCGGCGCTAGATGCGCCAG 740
Qy      541  ATGAGAAACGAGAAAGCTGCCGACGCGACGCTGATCTCCGCTACGCTGAGGACAGA 600
Db      741  ATGAGAAACGAGAAAGCTGCCGACGCGACGCTGATCTCCGCTACGCTGAGGACAGA 800
Qy      601  CTGGCCACCTCTACAGCGCTGTGTCTACTCTCCGCGTCCGACGAGGACACGCCCCTG 660
Db      801  CTGGCCACCTCTACAGCGCTGTGTCTACTCTCCGCGTCCGACGAGGACACGCCCCTG 860
Qy      661  GACCGGAGACACGCGGCGGCTGAGCGCCCTGGGACACAGGCTGTGTGGTGGCCACCTGTGC 720
Db      861  GACCGGAGACACGCGGCGGCTGAGCGCCCTGGGACACAGGCTGTGTGGTGGCCACCTGTGC 920
Qy      721  AGCGAGTTTGGGCTCTGAGACGCGACACTATCTGATCTCTGCGGGGACACGCTCATCTC 780
Db      921  AGCGAGTTTGGGCTCTGAGACGCGACACTATCTGATCTCTGCGGGGACACGCTCATCTC 980
Qy      781  TCGCAGGAGAAAGCCGTTGACGACACTACCTGGGGCTACTGCACTTTGTGAAGATTTTC 840
Db      981  TCGCAGGAGAAAGCCGTTGACGACACTACCTGGGGCTACTGCACTTTGTGAAGATTTTC 1040
Qy      841  TCCAAACTCTGTGGCTTCTTCAGACAGCTTTGTGACACCACTTCTACCGCTACATGAGAC 900
Db      1041  TCCAAACTCTGTGGCTTCTTCAGACAGCTTTGTGACACCACTTCTACCGCTACATGAGAC 1100
Qy      901  CAGAGTCTCCCGACGAGCTCCAAAGGCTGATGAAGAAAGCTGCTGCGGGGACCCGCGAC 960
Db      1101  CAGAGTCTCCCGACGAGCTCCAAAGGCTGATGAAGAAAGCTGCTGCGGGGACCCGCGAC 1160
Qy      961  TGTCTCCCGGACCAATGAGGAGTGCAGAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGC 1002
Db      1161  TGTCTCCCGGACCAATGAGGAGTGCAGAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGC 1202

```

RESULT 4

AA\$42859
ID AA\$42859 standard; cDNA; 1778 BP.

AA\$42859;

18-DEC-2001 (first entry)

Human G Protein-Coupled Receptor (GPCR) cDNA #54.

Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia; attention deficit disorder; anxiety; depression; bipolar disorder; ss; neurological disorder; Huntington's disease; dementia; obesity; anorexia; metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis; type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer; cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV; viral infection; immunostimulant; neuroleptic; nootropic; tranquilizer; antidepressant; anorectic; PCR primer; gene therapy.

Homo sapiens.

MO200162797-A2.

```

XX      30-AUG-2001.
PD      23-FEB-2001; 2001MO-US005676.
PF      23-FEB-2001; 2001MO-US005676.
XX      23-FEB-2001; 2000US-0184247P.
PR      23-FEB-2001; 2000US-0184303P.
PR      23-FEB-2001; 2000US-0184304P.
PR      23-FEB-2001; 2000US-0184305P.
PR      23-FEB-2001; 2000US-0184379P.
PR      02-MAR-2001; 2000US-0186437P.
PR      03-MAR-2001; 2000US-0186437P.
PR      09-MAR-2001; 2000US-0188064P.
PR      13-MAR-2001; 2000US-0188880P.
PR      03-APR-2001; 2000US-0194344P.
PR      23-JUN-2001; 2000US-0213861P.
PR      11-JUL-2001; 2000US-0217369P.
PR      11-JUL-2001; 2000US-0217370P.
PR      14-JUL-2001; 2000US-0218337P.
PR      20-JUL-2001; 2000US-0218492P.
XX      (PHMA ) PHARMACIA & UPJOHN CO.
XX      PA
PI      Vogel I G, Wood LS, Parodi LA, Lind P,
DR      WPI; 2001-570628/64.
DR      P-P8DB; AAU25607.
XX      New isolated nucleic acid encoding a new G-protein coupled receptor
PT      polypeptide for detecting receptor modulators that can treat mental
PT      disorders, such as schizophrenia, anxiety, depression, or obesity.
XX      Claim 4; Page 90-91; 279pp; English.
XX      Sequences AA$42806-AA$42926 represent cDNA molecules and PCR primers for
CC      cDNA molecules encoding human G-protein coupled receptor (GPCR)
CC      polypeptides. The protein and DNA sequences of the invention can be used
CC      to identify compounds which bind to GPCR polypeptides and in screening
CC      for compounds that modulate GPCR activity. By screening a human subject
CC      for the presence of mutations in GPCR DNA, a GPCR-related disorder or a
CC      genetic predisposition can be diagnosed. The sequences can also be used
CC      for treatment and prevention of mental disorders such as schizophrenia,
CC      attention deficit disorder, anxiety, depression, dementia and bipolar
CC      disorder, neurological disorders such as Huntington's disease,
CC      Parkinson's disease and Tourette's syndrome, metabolic disorders such as
CC      obesity, anorexia and type 2 diabetes, cardiovascular disorders such as
CC      thrombosis, myocardial infarction, cardiomyopathy and atherosclerosis,
CC      viral infections caused by HIV and cancers
XX      SQ
SQ      Sequence 1778 BP; 376 A; 541 C; 498 G; 363 T; 0 U; 0 Other;
Query Match 100.0%; Score 1002; DB 4; Length 1778;
Best Local Similarity 100.0%; Pred.No. 3.8e-194;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1  ATGTGAGCTGAGAGCTGTTCAACGCGACAGGCTGTGTGAGAGACTGCTTCCGCAAG 60
Db      19  ATGTGAGCTGAGAGCTGTTCAACGCGACAGGCTGTGTGAGAGACTGCTTCCGCAAG 78
Qy      61  GACTGACGCTGGGGCTGTCACTGTTGTGCTGCTGAGGCTGTGTGAGGCGTGCAGAG 120
Db      79  GACTGACGCTGGGGCTGTCACTGTTGTGCTGCTGAGGCTGTGTGAGGCGTGCAGAG 138
Qy      121  GGCTGTGTGTACAAGCGCTGTGTGTGCTGCGCAACTTACACAGAGCGACGATGACC 180
Db      139  GGCTGTGTGTACAAGCGCTGTGTGTGCTGCGCAACTTACACAGAGCGATGAGAC 198
Qy      181  ATGCGGAGCTGTACTTTGTCAACATGAGAGGAGGAGGCTGTGTAGGCGCTTGGCC 240
Db      199  ATGCGGAGCTGTACTTTGTCAACATGAGAGGAGGAGGCTGTGTAGGCGCTTGGCC 258
Qy      241  CTTGTGACCTGTGCGGCCCCCGAGCTCCGGTGGGCGCTGTGTGAGTGTGGGCGGAA 300

```

Db 259 CTGTGACACTGTGCTGGAGCCCCCGAGCTCCCGGAGCGCTGTGAGATGTGGCGGAGAA 318
 QY 301 GTCCAGTGTGCACTGAGATGCCCTTCAATGTGTCTCTCATCTGGTGGCAATGTACTCCACC 350
 Db 319 GTCCAGTGTGCACTGAGATGCCCTTCAATGTGTCTCTCATCTGGTGGCAATGTACTCCACC 378
 QY 361 GGCCTGTGAGCTTCGACCACTACATGAGAGTGCATGCGCGGAGCACTACATGAGCCAGC 420
 Db 379 GGCCTGTGAGCTTCGACCACTACATGAGAGTGCATGCGCGGAGCACTACATGAGCCAGC 438
 QY 421 GTGTACAAACACGCGGACAGTGTGCGGCTTGTGTGGGGTGGCGCGCTGTGACCAAGCTTC 480
 Db 439 GTGTACAAACACGCGGACAGTGTGCGGCTTGTGTGGGGTGGCGCGCTGTGACCAAGCTTC 498
 QY 481 TCCCTGCTGTCTTCTTACATCTGACGCAATGTGTCCACCGCGCGCTAAGTGTGGCCAG 540
 Db 499 TCCCTGCTGTCTTCTTACATCTGACGCAATGTGTCCACCGCGCGCTAAGTGTGGCCAG 558
 QY 541 ATGCAAAACGACAGAGCTGCGACGCGCAGCTGATGTTCAATCGGATACGTTGCGCAGCA 600
 Db 559 ATGCAAAACGACAGAGCTGCGACGCGCAGCTGATGTTCAATCGGATACGTTGCGCAGCA 618
 QY 601 CTGGCAACCTCTACGCGCTGTGTACTCTCCGCGCTCCGACGAGAGACACGCGCCCTG 660
 Db 619 CTGGCAACCTCTACGCGCTGTGTACTCTCCGCGCTCCGACGAGAGACACGCGCCCTG 678
 QY 661 GACCGGACACGCGGCGCGCTGTGAGCCTCTCGGACACAGGCTGTGTTGGTCCACCGTGTGC 720
 Db 679 GACCGGACACGCGGCGCGCTGTGAGCCTCTCGGACACAGGCTGTGTTGGTCCACCGTGTGC 738
 QY 721 AGCAGTTTGGGCTGTGAGAGCGACACTATCTGCTGCGGGGACACAGGTCATCATC 780
 Db 739 AGCAGTTTGGGCTGTGAGAGCGACACTATCTGCTGCGGGGACACAGGTCATCATC 798
 QY 781 TCGCAGGAGAAAGCCCGTGTGACGACACTACTCTGGGCTACTGTGACCTTGTGAAGATTTC 840
 Db 799 TCGCAGGAGAAAGCCCGTGTGACGACACTACTCTGGGCTACTGTGACCTTGTGAAGATTTC 858
 QY 841 TCCAACTCTGCGCTTCTTCCAGACAGTTTGTGACACACTTCTTACCGCTACATGAAAC 900
 Db 859 TCCAACTCTGCGCTTCTTCCAGACAGTTTGTGACACACTTCTTACCGCTACATGAAAC 918
 QY 901 CAGAGCTTCCCCGACAAAGCTCCAAAGGCTGATGAAAAAGCTGCGCGGAGACCGGAC 960
 Db 919 CAGAGCTTCCCCGACAAAGCTCCAAAGGCTGATGAAAAAGCTGCGCGGAGACCGGAC 978
 QY 961 TGCTCCCGGACCACTGGGGGTGACAGCAGTGTGCGCTAG 1002
 Db 979 TGCTCCCGGACCACTGGGGGTGACAGCAGTGTGCGCTAG 1020

RESULT 5
 AB242598 standard; DNA; 1785 BP.
 XX
 AC AB242598;
 XX
 DT 04-MAR-2003 (first entry)
 XX
 DE Human G protein-coupled receptor MGC7035 nucleotide seq ID NO.679.
 XX
 KM G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KM G protein-coupled receptor modulator; antibody; immune-related disease;
 KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KM immunological-related cell proliferative disease; autoimmune disease;
 KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KM osteoporosis; cardiovascular; inflammation; Crohn's disease; diabetes;
 KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KM mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 XX ulcer; gene; db.

OS Homo sapiens.
 XX
 PN MO200261087-A2.
 XX
 PD 08-AUG-2002.
 XX
 PP 19-DEC-2001; 2001WO-US050107.
 XX
 PR 19-DEC-2000; 2000US-0257144P.
 XX
 PA (LIFE-) LIFESPAN BIOSCIENCES INC.
 XX
 PI Burner GC, Roush CL, Brown JP;
 DR WPI; 2003-046718/04.
 DR P-PSDB; ABP81753.
 XX
 PT New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 XX
 PS Disclosure; Fig 1; 523pp; English.
 XX
 CC The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 1785 BP; 381 A; 539 C; 501 G; 364 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1002; DB 8; Length 1785;
 Best Local Similarity 100.0%; Pred. No. 3.8e-194;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTGAGCTGCAAGCTGTTCAACGACAGGCTGTGTGAGAGAGTCTGCTGCTCCAG 60
 Db 21 ATGTGAGCTGCAAGCTGTTCAACGACAGGCTGTGTGAGAGAGTCTGCTGCTCCAG 80
 QY 61 GACCTGACAGTGGGGGTGTCAGTGTGTGCTGCTGGGCTGTGGTGGGCGTGCAGAG 120
 Db 81 GACCTGACAGTGGGGGTGTCAGTGTGTGCTGCTGGGCTGTGGTGGGCGTGCAGAG 140
 QY 121 GGCCTGTGTTACAAAGCCTGTGCTGTGCTGACCACTTACACAGCAAGCCAGCATGACC 180
 Db 141 GGCCTGTGTTACAAAGCCTGTGCTGTGCTGACCACTTACACAGCAAGCCAGCATGACC 200
 QY 181 ATGCCGAGAGTGTACTTGTCAACATGGAGTGGCAGGCGCTGTGCTCAGCGCCCTGGCC 240
 Db 201 ATGCCGAGAGTGTACTTGTCAACATGGAGTGGCAGGCGCTGTGCTCAGCGCCCTGGCC 260
 QY 241 CCTGTGACACTGTGCGGCCCCCGAGCTCCGAGTGGGCGCTGTGAGTGTGGGCGGAGAA 300

Db CCTGGGACCTGCTGCGCCCGCCGAGCTCCGGTGGGGGCTGTGGAGTGGGGGGGAG 320
 QY GTCCACGGGGGACGACGAGATCCCTTAAATGTGCTCACTGTGGGCAATGATCCACC 360
 Db GTCCACGGGGGACGACGAGATCCCTTAAATGTGCTCACTGTGGGCAATGATCCACC 380
 QY GCCCTGTGAGCTTCGACCACTATGAGGTGCACTGGCGGGAGCTTACATGGCCAGC 420
 Db GCCCTGTGAGCTTCGACCACTATGAGGTGCACTGGCGGGAGCTTACATGGCCAGC 440
 QY GTGACAAACAGCGGCGACGCTGTGGGCTGTGGGGGTGGCGGCTGACCAAGCTTC 480
 Db GTGACAAACAGCGGCGACGCTGTGGGCTGTGGGGGTGGCGGCTGACCAAGCTTC 500
 QY TCCTCGCTGCTCTTCTCATCTGAGGCAATGTGCCACCGCGCGGCTAGAGTGGCCAG 540
 Db TCCTCGCTGCTCTTCTCATCTGAGGCAATGTGCCACCGCGCGGCTAGAGTGGCCAG 560
 QY ATGCAAGACGAGAAAGCTGCCAGCGCTGGTTCATCGGCTACGTGTGCCAGCA 600
 Db ATGCAAGACGAGAAAGCTGCCAGCGCTGGTTCATCGGCTACGTGTGCCAGCA 620
 QY CTGGCCACCTCTTAAACGGCTGTGTGCTACTCTCCCGGCTCCGAGGGAGAGCAAGCCCTTG 660
 Db CTGGCCACCTCTTAAACGGCTGTGTGCTACTCTCCCGGCTCCGAGGGAGAGCAAGCCCTTG 680
 QY GACCGGGACACGAGGCGCGCTGAGAGCCCTGGGACACAGAGGTGTGTGGTGCACCGGTGTC 720
 Db GACCGGGACACGAGGCGCGCTGAGAGCCCTGGGACACAGAGGTGTGTGGTGCACCGGTGTC 740
 QY AGCGAGTTGGGCTCTGAGACGCACTATCTGATCTCTGTGGGGGACACGGTCAATCATC 780
 Db AGCGAGTTGGGCTCTGAGACGCACTATCTGATCTCTGTGGGGGACACGGTCAATCATC 800
 QY TCGCAGAGGAAAGCCCGGAGACGCACTACCTGGGGCTTCTGACCTTTGTGAAGATTTC 840
 Db TCGCAGAGGAAAGCCCGGAGACGCACTACCTGGGGCTTCTGACCTTTGTGAAGATTTC 860
 QY TCCAACTCTGGCTCTTCTCCAGACGCTTGTGACACCACTTCTACCGCTACATGAAC 900
 Db TCCAACTCTGGCTCTTCTCCAGACGCTTGTGACACCACTTCTACCGCTACATGAAC 920
 QY CAGAGTTCCCGAGCAAGCTCCCAACGGCTGATGAAGAAAAGCTCCCTGCGGGAGACCGGAC 960
 Db CAGAGTTCCCGAGCAAGCTCCCAACGGCTGATGAAGAAAAGCTCCCTGCGGGAGACCGGAC 980
 QY TGCTCCCGGACCACTGGGGGTGACGAGGTGCTGGCGTAG 1022
 Db TGCTCCCGGACCACTGGGGGTGACGAGGTGCTGGCGTAG 1022

RESULT 6
 ADO30022
 ID ADO30022 standard; cDNA; 1785 BP.
 AC ADO30022;
 XX 29-JUL-2004 (first entry)
 DT
 XX
 DE Human GPCR PGR8 polynucleotide, SEQ ID NO:1124.
 KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thyroid disorder; thymus disorder; antiparkinsonian; autonomic;
 KW cytosolic; antiinflammatory; vasodilator; antianemic; antidiabetic;
 KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;

KW v1rnucleide, hepatotropic; antibacterial; antianemic; antiseborrheic;
 KW dermatological; antidiabetic; antihypertensive; anorectic;
 KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KW gene; ss.
 OS Homo sapiens.
 XX
 PN MO2004040000-A2.
 XX
 XX 13-MAY-2004.
 PD
 PF 09-SEP-2003; 2003WO-US028226.
 XX
 PR 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX
 PA (PRIM-) PRIMAL INC.
 PI Galenaris GA, Bergmann JB, Gragerov A, Hohmann J, Li F;
 PI Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
 XX
 DR WPI; 2004-390329/36.
 P-PSDB; ADO29606.
 XX
 PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 PS
 XX Claim 151; SEQ ID NO 1124; 542bp; English.
 XX
 CC The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancer). The present sequence represents a GPCR-encoding
 CC nucleic acid of the invention. Note: The full sequence data for this
 CC patent did not form part of the printed specification; those sequences
 CC not shown were obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1785 BP; 381 A; 539 C; 501 G; 364 T; 0 U; 0 Other;

Query Match 100.0%; Score 1002; DB 12; Length 1785;
 Best Local Similarity 100.0%; Pred. No. 3.8e-194;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGCGACGCTGTTCACCGGACAGGCTGTGTGAGAGAGTGTGCTGCTGCCAG 60
 Db 21 ATGTGAGCTGCGACGCTGTTCACCGGACAGGCTGTGTGAGAGAGTGTGCTGCTGCCAG 80

```

QY 61 GACCTGAGCTGGGGGCTGTCACTGTTGTGTGCTGGGCTGGTGGTGGGCGTGGCAAGT 120
DB 81 GACCTGAGCTGGGGGCTGTCACTGTTGTGTGCTGGGCTGGTGGTGGGCGTGGCAAGT 140
QY 121 GGGCTGTGTACACACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 141 GGGCTGTGTACACACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 200
QY 181 ATGCCGAGCTGTACTTTGTCAACATGGCAGTGGCAGGCGCTGGTGTCAAGCGCTGGCC 240
DB 201 ATGCCGAGCTGTACTTTGTCAACATGGCAGTGGCAGGCGCTGGTGTCAAGCGCTGGCC 260
QY 241 CCTGTGACCTGTGTGGGCGCGCGAGCTCCGGTGGGCGCTGTGTGTGTGTGTGTGTGTGT 300
DB 261 CCTGTGACCTGTGTGGGCGCGCGAGCTCCGGTGGGCGCTGTGTGTGTGTGTGTGTGTGT 320
QY 301 GTTCAGTGGCAGTGTGAGATCCCTTCAATGTGTCTCACTGTGTGTGTGTGTGTGTGTGT 360
DB 321 GTTCAGTGGCAGTGTGAGATCCCTTCAATGTGTCTCACTGTGTGTGTGTGTGTGTGTGT 380
QY 361 GCGCTGTGAGCTGTGACCACTATGAGCTGTGACCTGTGACCTGTGACCTGTGACCTGT 420
DB 381 GCGCTGTGAGCTGTGACCACTATGAGCTGTGACCTGTGACCTGTGACCTGTGACCTGT 440
QY 421 GTGTAAACACCGGGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB 441 GTGTAAACACCGGGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 500
QY 481 TCCTGCTGCTCTTCTACATCTGTGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
DB 501 TCCTGCTGCTCTTCTACATCTGTGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 560
QY 541 ATGCAGAACGACAGAGCTGTGACCGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
DB 561 ATGCAGAACGACAGAGCTGTGACCGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 620
QY 601 CTGGGCAACCTCTACACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
DB 621 CTGGGCAACCTCTACACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 680
QY 661 GACCGGAGACACGGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
DB 681 GACCGGAGACACGGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 740
QY 721 AGCGAGTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 741 AGCGAGTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 800
QY 781 TGGCGAGGAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 801 TGGCGAGGAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 860
QY 841 TCCAAACTCTGTGCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
DB 861 TCCAAACTCTGTGCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 920
QY 901 CAGAGCTTCCCGACAGCTTCCAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
DB 921 CAGAGCTTCCCGACAGCTTCCAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 980
QY 961 TGGTCCCCCGGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1002
DB 981 TGGTCCCCCGGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022

```

```

RESULT 7
ACCT0841 standard; mRNA; 1926 BP.
XX
ACCT0841;
XX
DT 20-NOV-2003 (first entry)
XX

```

```

DE Human G-protein coupled receptor 901 mRNA.
XX Human; anorectic; antidiabetic; antilipemic; hypothalamus;
KW G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
KW clobophobla; anorexia nervosa; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 154..1155
PT /tag= a
PT /product= "G-protein coupled receptor 901"
XX
PN MO200303936-A1.
XX
PD 17-APR-2003.
XX
PF 02-OCT-2002; 2002WO-JP010250.
XX
PR 02-OCT-2001; 2001JP-00306872.
XX
PA (SUMU ) SUMITOMO PHARM CO LTD.
XX
PI Suguru E, Tsuchida A, Yamanaka M, Tajji M;
XX
DR MPI; 2003-354886/33.
XX
DR P-PSDB; ABR56303.
XX
PT Inhibitors of expression or activity of G-protein coupled receptor 901
XX for treatment of lifestyle-related diseases and clobophobia.
XX
PS Claim 2; Page 62-64; 91pp; Japanese.
XX
CC The present invention relates to novel remedies for the treatment of
CC diseases containing an active component an inhibitor of the expression
CC or activity of hypothalamus-expressed G-protein coupled receptor 901 and
CC for treatment of clobophobia containing an active component a
CC potentiator of the expression or activity of G-protein coupled receptor
CC 901. The diseases which can be treated include obesity, diabetes and
CC hyperlipaemia, and clobophobia (anorexia nervosa). The present human
CC sequence is the mRNA for G-protein coupled receptor 901
XX
SQ Sequence 1926 BP; 407 A; 590 C; 546 G; 0 T; 383 U; 0 Other;
Query Match 100.0%; Score 1002; DB 8; Length 1926;
Best Local Similarity 80.7%; Pred. No. 3.8e-194;
Matches 809; Conservative 193; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGAGCTGTGAGCTGTTCACAGGACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
DB 154 ATGTGAGCTGTGAGCTGTTCACAGGACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 213
QY 61 GACCTGAGCTGGGCTGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
DB 214 GACCTGAGCTGGGCTGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 273
QY 121 GGCCTGTGTACACACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 274 GGCCTGTGTACACACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 333
QY 181 ATGCCGAGCTGTACTTTGTCAACATGGCAGTGGCAGGCGCTGGTGTCAAGCGCTGGCC 240
DB 334 ATGCCGAGCTGTACTTTGTCAACATGGCAGTGGCAGGCGCTGGTGTCAAGCGCTGGCC 393
QY 241 CCTGTGACCTGTGTGGGCGCGCGAGCTCCGGTGGGCGCTGTGTGTGTGTGTGTGTGTGT 300
DB 394 CCTGTGACCTGTGTGGGCGCGCGAGCTCCGGTGGGCGCTGTGTGTGTGTGTGTGTGTGT 453
QY 301 GTTCAGTGGCAGTGTGAGATCCCTTCAATGTGTCTCACTGTGTGTGTGTGTGTGTGTGT 360
DB 454 GTTCAGTGGCAGTGTGAGATCCCTTCAATGTGTCTCACTGTGTGTGTGTGTGTGTGTGT 513
QY 361 GCGCTGTGAGCTGTGACCACTATGAGCTGTGACCTGTGACCTGTGACCTGTGACCTGT 420

```


[illegible]

Result	ID	Accession	Gene	Protein	Function
RESULT 8	ACAG1176	ACAG1176	CDNA, 1928 BP.	Human, ss; gene; G-protein coupled receptor; gene therapy; pneumonia; immune disorder; Crohn's disease; Grave's disease; respiratory disorder; diphtheria; haematopoietic disorder; leukaemia; systemic sclerosis; prostatic disorder; benign prostate hypertrophy; tumour; breast disease; acute mastitis; Paget's disease; muscular disorder; Thabomvosa; atrophy; neurological disorder; cerebral odema; Parkinson's disease; atrophy; blood vessel disorder; atherosclerosis; testicular disease; syphilis; epididymal disease; thyroid disease; hyperthyroidism; cretinism; AIDS; kidney disorder; cystic renal dysplasia; glomerulonephritis; cardiovascular disease; heart failure; pericarditis; pancreatic disease; pancreatitis; diabetes mellitus; thymus disease; thymic hypoplasia; Hodgkin disease; spleen disease; splenomegaly; Gaucher disease; liver disease; hepatic failure; alcoholic liver disease; T-cell disorder; systemic lupus erythematosus; B-cell disorder; Burkitt lymphoma; multiple myeloma; platelet disorder; thrombocytopenia; haemolytic-uraemic syndrome.	Human sapiens.

Key	Location/Qualifiers
FT CDS	160..1161
FT	/*tag= a
FT	/product= "GPCR 23992"
FT	/note= "G-protein coupled receptor"
XX	US2003017539-A1.
XX	23-JAN-2003.
XX	07-JUN-2002; 2002US-00165844.
PR	02-JUN-1998; 98US-00088857.
PR	02-SEP-1998; 98US-00145745.
PR	21-JAN-1999; 99US-00234923.
PR	02-JUN-1999; 99US-00324465.
PR	28-JUN-1999; 99US-00340880.
PR	26-AUG-1999; 99US-00383745.
PR	16-DEC-1999; 99US-00464665.
PR	18-DEC-2000; 2000US-00741783.
XX	(MILL-) MILLENNIUM PHARM INC.
PA	Gluckemann MA, Hodge NR, Hunter JJ, Rudolph-Owen LA;
PI	SIlos-Santiago I, Weich NS;
XX	WPI; 2003-401672/38.
DR	P-PSDB; AB009900.
XX	New nucleic acid molecule encoding a G-protein coupled receptor
PT	polypeptide, e.g. 2871, 14926, 17723 or 23992, useful for diagnosing
PT	cardiovascular disorders.
PT	cardiovascular disorders.
XX	Claim 2, Fig 21, 1499p; English.
XX	The invention relates to an isolated G-protein coupled receptor nucleic
XX	acid molecule encoding the G-protein coupled receptors 2871, 14926, 17722
CC	and 23992. The nucleic acid molecule is useful in monitoring, diagnosing
CC	and treating immune disorders (e.g. Crohn's disease and Grave's disease)
CC	respiratory disorders (e.g. pneumonia or diptheria); haematopoietic
CC	disorders (e.g. leukaemia and systemic sclerosis); prostatic disorders
CC	(e.g. benign prostate hypertrophy and tumours); breast diseases (e.g.
CC	acute mastitis and Paget's disease); muscular disorders (e.g.
CC	rhabdomyosarcoma); neurological disorders (e.g. cerebral oedema and
CC	Parkinson's disease); disorders involving blood vessels (e.g.
CC	atherosclerosis); diseases involving testis and epididymis (e.g.
CC	and syphilis); thyroid diseases (e.g. hyperthyroidism and cretinism);
CC	kidney disorders (e.g. cystic renal dysplasia and glomerulonephritis);
CC	cardiovascular diseases (e.g. cystic heart failure and pericarditis); pancreatic
CC	diseases (e.g. pancreatitis and diabetes mellitus); diseases involving
CC	the thymus (e.g. thymic hypoplasia and Hodgkin disease); diseases
CC	involving the spleen (e.g. splenomegaly and Gaucher disease); diseases
CC	involving the liver (e.g. hepatic failure and alcoholic liver disease);
CC	disorders (e.g. systemic lupus erythematosus and AIDS); B-cell
CC	(e.g. chronic lymphoma and multiple myeloma); platelet disorders
CC	(e.g. thrombocytopenia and haemolytic-uremic syndrome). The nucleic
CC	acids may also be used in chromosome mapping, tissue typing,
CC	pharmacogenomics and forensic biology, and as surrogate markers. The
CC	present sequence represents cDNA encoding the human G-protein coupled
CC	receptor 23992
XX	Sequence 1928 BP, 408 A, 588 C, 547 G, 385 T, 0 U, 0 Other;
XX	Query Match 100.0%; Score 1002; DB 8; Length 1928;
XX	Best Local Similarity 100.0%; Pred. No. 3 8e-194;
XX	Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX	1 ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 60
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 120
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 180
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 240
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 300
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 360
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 420
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 480
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 540
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 600
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 660
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 720
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 780
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 840
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 900
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 960
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1020
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1080
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1140
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1200
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1260
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1320
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1380
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1440
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1500
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1560
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1620
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1680
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1740
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1800
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1860
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1920
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 1980
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 2040
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 2100
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 2160
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 2220
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 2280
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 2340
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 2400
XX	ATGTGAGCTGCAAGCTGTTCAACGGACAGAGGCTGTGTGAGAGAGCTGCCTGCTGCAG 2460


```

Db      220 GACCTGACGCTGGGCGCTGTACCTGTGTGCTGCTGGGCTGGTGGGCGTGGCCAGT 279
Qy      121 GGCCTGTGTACACACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db      280 GGCCTGTGTACACACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 339
Qy      181 ATGCCGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db      340 ATGCCGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 399
Qy      241 CTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Db      400 CTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 459
Qy      301 GTCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db      460 GTCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519
Qy      361 GGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Db      520 GGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579
Qy      421 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db      580 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 639
Qy      481 TCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db      640 TCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 699
Qy      541 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db      700 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 759
Qy      601 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db      760 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 819
Qy      661 GACCGGACACGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db      820 GACCGGACACGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 879
Qy      721 AGCGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db      880 AGCGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 939
Qy      781 TCGGAGGAGGAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db      940 TCGGAGGAGGAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 999
Qy      841 TCCAAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db      1000 TCCAAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1059
Qy      901 CAGAAGTTTCCCGACGAGCTTCCAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Db      1060 CAGAAGTTTCCCGACGAGCTTCCAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1119
Qy      961 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1002
Db      1120 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1161

```

```

RESULT 9
AAZ90175
ID      AAZ90175 standard; cDNA; 1614 BP.
XX      AC      AAZ90175;
XX      DT      19-MAY-2000 (first entry)
XX      XX      Human chemokine receptor nucleotide sequence.

```

```

XX      KW      Chemokine receptor; interleukin-8 compound inhibitor; chromosome 7p22;
KW      inflammation; wound healing; neutropenia; myeloid leukaemia; tumour; ss;
KW      toxin delivery; hypermegakaryocytopenic disease; polycythaemia vera.
XX      OS      Homo sapiens.
XX      FT      CDS
XX      Key      Location/Qualifiers
XX      CDS      129..1127
XX      FT      /tag=a
XX      FT      /product="Chemokine receptor"
XX      FT      /transl_except=(Pos:159..161, aa:Xaa)
XX      FT      /transl_except=(Pos:177..179, aa:Xaa)
XX      FT      /note="Xaa = Unknown"
XX      PN      WO200000515-A2.
XX      PD      06-JAN-2000.
XX      PE      29-JUN-1999; 99WO-US012829.
XX      PR      29-JUN-1998; 98US-00106800.
XX      PR      22-JAN-1999; 99US-00236166.
XX      PA      (HYSE-) HYSBQ INC.
XX      DR      MPI; 2000-170907/15.
XX      DR      P-PSDB; AAY78856.
XX      PT      New nucleic acid encoding chemokine receptor, useful for diagnosis and
XX      PT      treatment of e.g. neutropenia, inflammation and leukemia.
XX      PS      Claim 1; Fig 1; 138pp; English.
XX      CC      This sequence represents a human chemokine receptor coding sequence. The
XX      CC      sequence is derived from a sequence isolated from a human foetal liver-
XX      CC      spleen cDNA library. The chemokine receptor (see AAY78856) encoded by the
XX      CC      extended nucleotide sequence inhibits the activity of interleukin-8-type
XX      CC      compounds through competition for cell binding sites. The chemokine
XX      CC      receptor gene is located on the short arm of chromosome 7 at 7p22. The
XX      CC      polynucleotide encoding the chemokine receptor is useful as a
XX      CC      hybridization probe or a PCR primer, the nucleotide sequence may also be
XX      CC      used for chromosome/gene mapping or in the recombinant production of
XX      CC      polypeptides and the production of antisense or triplex-forming molecules
XX      CC      for the control of gene expression. The chemokine receptor polypeptides
XX      CC      are used to raise specific antibodies, also for purification, detection
XX      CC      or modulation of interleukin-8-type chemokines (for diagnosis or
XX      CC      prognosis, or monitoring chemokine recruitment at a site of infection or
XX      CC      inflammation). The protein sequence can also be used as molecular weight
XX      CC      markers or food supplements, and to screen compound libraries for
XX      CC      specific binding agents, potential agonists or antagonists. Antibodies
XX      CC      raised against the chemokine receptor polypeptide sequence are used to
XX      CC      detect or purify the polypeptide, also for the diagnosis and treatment of
XX      CC      activated or inflamed cells or tissues, and to promote the healing of
XX      CC      wounds. The polypeptide and antibodies are also used to prevent
XX      CC      neutropenia (associated with chemotherapy or radiation treatment to
XX      CC      protect myeloid precursors), inflammation or other immune responses; also
XX      CC      conditions associated with hyperproliferation of progenitor cells (e.g.
XX      CC      some myelogenous leukaemias, polycythaemia vera and
XX      CC      hypermegakaryocytopenic diseases). The antibodies are potentially
XX      CC      useful therapeutically, e.g. to carry toxins to tumour cells
XX      SQ      Sequence 1614 BP, 307 A, 509 C, 468 G, 326 T, 0 U, 4 Other;
XX
XX      Query Match      99.9%; Score 1000.8; DB 3; Length 1614;
XX      Best Local Similarity 99.7%; Pred. No. 6.6e-194;
XX      Matches 999; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```


Db	199	GACCTGCACTGGGGGTGTCACTGTTGTGGCTGTGGGCTGGTGGTGGGCGTGGCAGT	258
Oy	121	GGCCTGTGTACAACGCGCCCTGTGCTGTGGCTGTGGCCACTACACAGCAAGCCACATGAC	180
Db	259	GGCCTGTGTACAACGCGCCCTGTGCTGTGGCTGTGGCCACTACACAGCAAGCCACATGAC	31.8
Oy	181	ATGCCGGAGGTGTACTTTTGTCAACATGGGAGTGGGAGGCTTGGTGTCAAGCGCTTGGCC	240
Db	319	ATGCCGGAGGTGTACTTTTGTCAACATGGGAGTGGGAGGCTTGGTGTCAAGCGCTTGGCC	37.8
Oy	241	CCTGTGACCTGTGCGGCCCCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGGA	300
Db	379	CCTGTGACCTGTGCGGCCCCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGGA	438
Oy	301	GTCACGTGGCACTGCAGATCCCTTTCAATGTGTCTCACTGGTGGCCATGTACTCAC	360
Db	439	GTCACGTGGCACTGCAGATCCCTTTCAATGTGTCTCACTGGTGGCCATGTACTCAC	498
Oy	351	GGCCTGTGTGAGCTTGCACCACTACATGAGCGTGCACCTGCCGGACCTACATGGCCAGC	420
Db	499	GGCCTGTGTGAGCTTGCACCACTACATGAGCGTGCACCTGCCGGACCTACATGGCCAGC	558
Oy	421	GTGTACAAACAGCGGAGGTGTGGCGGCTTGTTGGGGTGGGCGGCTGCACAGCTTC	480
Db	559	GTGTACAAACAGCGGAGGTGTGGCGGCTTGTTGGGGTGGGCGGCTGCACAGCTTC	618
Oy	481	TCTGTGCTGTCTTCTTACACTGTGCAGCCATGTGTCCACCAGCGCTAGAGTGGCCAG	540
Db	619	TCTGTGCTGTCTTCTTACACTGTGCAGCCATGTGTCCACCAGCGCTAGAGTGGCCAG	678
Oy	541	ATGCAAGAACGAGAAAGCTGCCAGCCGACAGGCTGTGTTCATCGGCTACGTGTGCCAGCA	600
Db	679	ATGCAAGAACGAGAAAGCTGCCAGCCGACAGGCTGTGTTCATCGGCTACGTGTGCCAGCA	738
Oy	601	CTGGCACCCCTCTAAGCGCTGTGGTCACTCCCGGCTCCGAGGGAGGACAGGCCCTTG	660
Db	739	CTGGCACCCCTCTAAGCGCTGTGGTCACTCCCGGCTCCGAGGGAGGACAGGCCCTTG	798
Oy	661	GACCGGGACACGCGGCGGCTGTGAGCCCTCGGACACAGGCTGTGTGGCCACCGTGTGC	720
Db	799	GACCGGGACACGCGGCGGCTGTGAGCCCTCGGACACAGGCTGTGTGGCCACCGTGTGC	858
Oy	721	ACGCAAGTTTGGGCTCTGAGCGCCACACTATCTGATCTGTGGGGGACACAGGTCATCATC	780
Db	859	ACGCAAGTTTGGGCTCTGAGCGCCACACTATCTGATCTGTGGGGGACACAGGTCATCATC	918
Oy	781	TGCGAAGGGAAGCCCGTGTGACGCAACTACCTGGGGCTACTGACTTTGTGAAGATTTC	840
Db	919	TGCGAAGGGAAGCCCGTGTGACGCAACTACCTGGGGCTACTGACTTTGTGAAGATTTC	978
Oy	841	TCCAAACTCTGTGGCTTCTCCAGCAGCTTGTGTACACCACTTCTTACCGCTACATGAAC	900
Db	979	TCCAAACTCTGTGGCTTCTCCAGCAGCTTGTGTACACCACTTCTTACCGCTACATGAAC	103
Oy	901	CAGAGCTTCCCGAGCAAGCTCCCAAGGCTGTATGAATAAAGCTGCCCTTGCAGGAGCCGAC	960
Db	1039	CAGAGCTTCCCGAGCAAGCTCCCAAGGCTGTATGAATAAAGCTGCCCTTGCAGGAGCCGAC	109
Oy	961	TGCTTCCCGAGCAACTATGGGGGTGTGACGAGGTGTCTGGCTAG 1002	
Db	1099	TGCTTCCCGAGCAACTATGGGGGTGTGACGAGGTGTCTGGCTAG 1140	
RESULT 11			
ABX73053			
ID	ABX73053 standard; cDNA; 1878 BP.		
XX	ABX73053;		
AC			
XX	14-MAR-2003 (first entry)		
DT			
XX	Human G-protein coupled receptor cDNA #3, Incyte clone 2214673CB1.		

XX	Human; sg; gene; G-protein coupled receptor; GPCR; neoplastic disorder;
KM	neurological disorder; immune disorder; cystostatic; pancreatic cancer;
KM	follicular carcinoma of the thyroid; leiomyoma of the uterus; epilepsy;
KW	interstitial nephritis; cancer.
XX	
OS	Homo sapiens.
XX	
FN	US2002106655-A1.
PD	08-AUG-2002.
XX	
PR	28-JUN-2001; 2001US-00895666.
XX	
PR	17-SEP-1998; 98US-00156513.
XX	
PA	(BAND/) BANDMAN O.
PA	(LALP/) LAL P G.
PA	(TANG/) TANG Y T.
PA	(BADG/) BAUGHN M R.
PI	Bandman O, Lal PG, Tang YT, Baughn MR;
XX	
DR	WPI; 2002-697866/75.
XX	P-PSDB; ABUS4867.
PT	New cDNAs encoding G protein coupled receptors are useful for the
PT	diagnosis, prognosis, treatment and evaluation of therapies for
PT	neoplastic, neurological and immune disorders.
XX	
PS	Claim 2; Page 30-31; 61pp; English.
XX	
CC	The invention relates to an isolated cDNA encoding G-protein coupled
CC	receptor (GPCR) appearing as ABUS4865-ABUS4870. Also included are
CC	fragments of the cDNAs, species variants having at least 75% identity to
CC	the cDNAs, vectors comprising the cDNAs, a host cells comprising the
CC	above vectors, producing a protein (comprising culturing the above host
CC	cell under expression conditions and recovering the protein), using a
CC	cDNA to detect expression of a nucleic acid in a sample or to screen for
CC	compounds or molecules which bind to the cDNAs, using the GPCR proteins
CC	to screen compounds or molecules for ligands, using a GPCR protein to
CC	prepare and purify antibodies, an anti-GPCR antibody and using the
CC	antibody to detect expression of a GPCR protein in a sample and is
CC	diagnostic of cancer. The invention is useful for the diagnosis,
CC	prognosis, treatment and evaluation of therapies for neoplastic,
CC	neurological and immune disorders, particularly follicular carcinoma of
CC	the thyroid, leiomyoma of the uterus, pancreatic cancer, epilepsy,
CC	interstitial nephritis and immune response as a complication of cancer.
CC	The present sequence encodes a human GPCR protein of the invention
XX	
SQ	Sequence 1878 BP; 382 A; 585 C; 533 G; 378 T; 0 U; 0 Other;
XX	
Query Match	99.8%; Score 1000.4; DB 6; Length 1878;
Best Local Similarity	99.9%; Pred. No. 8.1e-194;
Matches 1001; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
0Y	1 ATGTGAGAGCTGCACACTGGTTCAACGCGACAGAGGCTGTGTGAGAGAGCTGCTGCGCAG 60
Db	139 ATGTGAGAGCTGCACACTGGTTCAACGCGACAGAGGCTGTGTGAGAGAGCTGCTGCGCAG 198
0Y	61 GACCTGCAGCTGGGGGCTGTCACTGTGTTCGTGCTGGGCTGTGTGTGGGCGTGCAGTG 120
Db	199 GACCTGCAGCTGGGGGCTGTCACTGTGTTCGTGCTGGGCTGTGTGTGGGCGTGCAGTG 258
0Y	121 GGCGTGTGCTACCAACGGCCGTGTGTGTCTGGGCAACTTACACAAAGCCAGCATGACC 180
Db	259 GGCGTGTGCTACCAACGGCCGTGTGTGTCTGGGCAACTTACACAAAGCCAGCATGACC 318
0Y	181 ATGCGGAGCGGTACTTTGTTCACATGAGCAGTGGCAGGCTGTGCTCAGCGCCCTGGCC 240
Db	319 ATGCGGAGCGGTACTTTGTTCACATGAGCAGTGGCAGGCTGTGCTCAGCGCCCTGGCC 378
0Y	241 CTGTGTGACCTGTGCGGCCCCCGAGACTCCCGGTGGGCGCTGTGAGTGTGGCGGCGAA 300

```

Db      379 CCGTGTGACACCTGCTGGGCCCCCGAGCTCCGCTGGGCGCTGTGAGTGTGGAGCGGAGAA 438
QY      301 GTCCACAGTGGCACTGAGATCCCTTTCAATGTGTCTTCACTGTGTGGCCATGTACTCCACC 360
Db      439 GTCCACAGTGGCACTGAGATCCCTTTCAATGTGTCTTCACTGTGTGGCCATGTACTCCACC 498
QY      361 GCGCTGTGTAGAGCTGACCACTACATGAGCGGTGACGCGCGGAGACCTACATATGGCCAGC 420
Db      499 GCGCTGTGTAGAGCTGACCACTACATGAGCGGTGACGCGCGGAGACCTACATATGGCCAGC 558
QY      421 GTGTACACAGCGGCGCACTGTGTGGGCTTGTGTGGGTTGGCGGCTGTGACCAAGCTTC 480
Db      559 GTGTACACAGCGGCGCACTGTGTGGGCTTGTGTGGGTTGGCGGCTGTGACCAAGCTTC 618
QY      481 TCCTGTGTGTCTTTTACATCTGACGAGCATGTGTCCACCGCGCGGCTTGAATGTGGCCAG 540
Db      619 TCCTGTGTGTCTTTTACATCTGACGAGCATGTGTCCACCGCGCGGCTTGAATGTGGCCAG 678
QY      541 ATGCAGAAAGCAAGAGCTGCCAGCCACGCTGTGTTCATCGGCTACATGTGTGTCCAGCA 600
Db      679 ATGCAGAAAGCAAGAGCTGCCAGCCACGCTGTGTTCATCGGCTACATGTGTGTCCAGCA 738
QY      601 CTGGCCACACCTCTACGCGGTGTGTGTCTACTCTCCGCGTCCGACAGGAGAGACACAGCCCTG 660
Db      739 CTGGCCACACCTCTACGCGGTGTGTGTCTACTCTCCGCGTCCGACAGGAGAGACACAGCCCTG 798
QY      661 GACCGGGAGACAGGCGCGGCTGTGAGCCTCTCGGACACAGGAGTGTGGTGGTGGACACCGTGTG 720
Db      799 GACCGGGAGACAGGCGCGGCTGTGAGCCTCTCGGACACAGGAGTGTGGTGGTGGACACCGTGTG 858
QY      721 ACCGAGTTTGGGCTCTCTGAGAGCCACACTATCTGTCTGTGGGAGACACAGGTCATCATC 780
Db      859 ACCGAGTTTGGGCTCTCTGAGAGCCACACTATCTGTCTGTGGGAGACACAGGTCATCATC 918
QY      781 TCGGAGAGAGAGCCCGTGTGAGCAGCACTACTCTGGGCTCTGTGACTTTGTGAAGATTTC 840
Db      919 TCGGAGAGAGAGCCCGTGTGAGCAGCACTACTCTGGGCTCTGTGACTTTGTGAAGATTTC 978
QY      841 TCCAAACTCTGTGGCTTTCTCAGACAGCTTTGTGACACCACTTCTACCGCTACATGAAAC 900
Db      979 TCCAAACTCTGTGGCTTTCTCAGACAGCTTTGTGACACCACTTCTACCGCTACATGAAAC 1038
QY      901 CAGAGCTTCCCGCAGAGCTCCACAGGCGGTGATGAAAAAGCTCCCTCGCGGGAGACCGGAGC 960
Db      1039 CAGAGCTTCCCGCAGAGCTCCACAGGCGGTGATGAAAAAGCTCCCTCGCGGGAGACCGGAGC 1098
QY      961 TGTCTCCCGAGACCACTGTGGGTTGACAGAGTGTGGCGTAG 1002
Db      1099 TGTCTCCCGAGACCACTGTGGGTTGACAGAGTGTGGCGTAG 1140

```

RESULT 12
AAH99600 standard, cDNA: 1893 BP.

AAH99600;

16-OCT-2001 (first entry)

Human protein encoding cDNA sequence SEQ ID NO:435.

Human; cancer; ulcer; HIV infection; human immunodeficiency virus; anti-inflamatory; antirheumatic; antidiabetic; immunosuppressive; anti-bacterial; endocrine; cardiac; central nervous system; vitruide; anti-HIV; fungicide; antimitogen; cardiovascular; antiaiemic; anemia; antiagregant; haemostatic; vulnery; antilucer; osteopatic; eczema; dermatological; antiallergic; antiaiemic; antidiabetic; cyostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antiaiemic therapy; vaccine; inflammation; antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity; genetic disease; haematopoietic disorder; platelet disorder; asthma;

```

KW      thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW      allergic rhinitis; diabetes; multiple sclerosis; depression;
KW      Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW      neurological disorder; ss.
XX
OS      Homo sapiens.
XX
PN      WO200153455-A2.
XX
PD      26-JUL-2001.
XX
PF      22-DEC-2000; 2000WO-US035017.
XX
PR      23-DEC-1999; 99US-00471275.
PR      21-JAN-2000; 2000US-00488725.
PR      25-APR-2000; 2000US-00552317.
XX
PA      (HYSB-) HYSBQ INC.
XX
PI      Tang YT, Liu C, Dermanac RT;
XX
DR      WPI; 2001-457603/49.
XX
PT      P-PsDB; AMM25659.
XX
PS      Isolated human polynucleotides encoding polypeptides, useful for the
XX      treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX      Claim 1; Page 522-523; 1217p; English.
XX
XX      AAH99166 to AAH99904 encode the human proteins given in AMM25725 to
XX      AMM25963. The proteins can have activities based on the tissues and cells
XX      they are expressed in, such as: anti-inflamatory; antirheumatic;
XX      antidiabetic; immunosuppressive; antibacterial; endocrine; cardiac;
XX      central nervous system; vitruide; anti-HIV; fungicide; antilucer;
XX      cardiovascular; antiaiemic; antiagregant; haemostatic; vulnery;
XX      antilucer; osteopatic; dermatological; antiallergic; antiaiemic;
XX      antidiabetic; cyostatic; neuroprotective; antidepressant; nootropic;
XX      antiparkinsonian; and immunostimulant. The proteins and polynucleotides
XX      encoding them can be used in gene therapy, antiaiemic therapy and vaccine
XX      production. The proteins and polynucleotides are useful for screening for
XX      agonists or antagonists of a protein and for the treatment and diagnosis
XX      of disorders associated with the activity of a protein e.g. inflammation,
XX      rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
XX      infections, autoimmunity, genetic diseases, haematopoietic disorders,
XX      anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
XX      osteoporosis, severe combined immunodeficiency, eczema, allergic
XX      rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
XX      Alzheimer's disease, Parkinson's disease, neurodegenerative and
XX      neurological disorders
XX
SQ      Sequence 1893 BP; 395 A; 583 C; 533 G; 382 T; 0 U; 0 Other;

```

Query Match 99.8%; Score 1000.4; DB 4; Length 1893;
Best Local Similarity 99.9%; Pred. No. 8.1e-194;
Matches 1001; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 ATGTGAGTGTGAGCTGTGTTCAAGCGCACAGGCGTGTGTGAGAGAGCTGCTGCGCCAG 60
Db      138 ATGTGAGTGTGAGCTGTGTTCAAGCGCACAGGCGTGTGTGAGAGAGCTGCTGCGCCAG 197
QY      61 GACCTGAGCTGGGCGTGTGCTGTGTGCGTGGGCGGTGTGTGGGCGGTGTGCGCAAGT 120
Db      198 GACCTGAGCTGGGCGTGTGCTGTGTGCGTGGGCGGTGTGTGGGCGGTGTGCGCAAGT 257
QY      121 GCGCTGTGTGTACAGCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db      258 GCGCTGTGTGTGTACAGCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 317
QY      181 ATGCGGAGAGTGTGCTTTGTGACATGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db      318 ATGCGGAGAGTGTGCTTTGTGACATGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 377

```

QY 241 CTTGTGACCTGCTCGGCCCCCGAGCTCCCGGTGGCGCTGTGAGTGTGGCGGAA 300
DB 378 CTTGTGACCTGCTCGGCCCCCGAGCTCCCGGTGGCGCTGTGAGTGTGGCGGAA 437
QY 301 GTCCACGTGGCACTGCAAGATCCCTTCAATGTGTCTTCACTGTGGCCATGTATCCACC 360
DB 438 GTCCACGTGGCACTGCAAGATCCCTTCAATGTGTCTTCACTGTGGCCATGTATCCACC 497
QY 361 GCGCTGTGAGCTTCCACCACTACATGAGCGTGGCACTGCGCGGCACTACATGTGGCAGC 420
DB 498 GCGCTGTGAGCTTCCACCACTACATGAGCGTGGCACTGCGCGGCACTACATGTGGCAGC 557
QY 421 GTGTAACAACAGCGGCACTGTGTGGCGCTGTGTGGGTGGCGCGCTGTGACCAAGCTTC 480
DB 558 GTGTAACAACAGCGGCACTGTGTGGCGCTGTGTGGGTGGCGCGCTGTGACCAAGCTTC 617
QY 481 TCTCTGCTGCTTTTCAATCTGAGCCATGTGTCCACCGCGGCTTGAAGTGGCCAG 540
DB 618 TCTCTGCTGCTTTTCAATCTGAGCCATGTGTCCACCGCGGCTTGAAGTGGCCAG 677
QY 541 ATGCAAAACGCAAAAGTGGCGGAGCGCAAGCTGTGTTCATCGGCTTACGTGTGCCAGCA 600
DB 678 ATGCAAAACGCAAAAGTGGCGGAGCGCAAGCTGTGTTCATCGGCTTACGTGTGCCAGCA 737
QY 601 CTGGCCACCTCTACGCGCTGTGTGTCTACTCTCCGCGTGGCGAGGAGGAGCAAGCCCTTG 660
DB 738 CTGGCCACCTCTACGCGCTGTGTGTCTACTCTCCGCGTGGCGAGGAGGAGCAAGCCCTTG 797
QY 661 GACCGGAGCAAGCGCGGCTGTGAGCGCTTCCGCAACAAGGCTGTGTGGCCACCGTGTGC 720
DB 798 GACCGGAGCAAGCGCGGCTGTGAGCGCTTCCGCAACAAGGCTGTGTGGCCACCGTGTGC 857
QY 721 ACGCAATTTGGGCTCTGGAAGCCACATATCTGTCTGTGGGCGACAGCTCATATC 780
DB 858 ACGCAATTTGGGCTCTGGAAGCCACATATCTGTCTGTGGGCGACAGCTCATATC 917
QY 781 TCGGAGGAGGAGCGCGTGGAGGAGCACTACCTGGGCTTGTGAAAGGATTC 840
DB 918 TCGGAGGAGGAGCGCGTGGAGGAGCACTACCTGGGCTTGTGAAAGGATTC 977
QY 841 TCCAACTCTGAGCTTCTCTCAGAGCTTGTGACCACTTCTTACCGCTACATGAC 900
DB 978 TCCAACTCTGAGCTTCTCTCAGAGCTTGTGACCACTTCTTACCGCTACATGAC 1037
QY 901 CAGAAGTTCCTCCGCAAGCTTCCAAAGCTGTGATGAAAAAGCTTCCGCGGAGCCGAGC 960
DB 1038 CAGAAGTTCCTCCGCAAGCTTCCAAAGCTGTGATGAAAAAGCTTCCGCGGAGCCGAGC 1097
QY 961 TGTCCCGGAGCAATGGGGGTGAGAGAGTGTGGCGTAG 1002
DB 1098 TGTCCCGGAGCAATGGGGGTGAGAGAGTGTGGCGTAG 1139

RESULT 13
ACCT0860
ID ACC70860 standard; DNA; 2216 BP.
XX
ACCT0860;
XX
DT 20-NOV-2003 (first entry)
XX
DE Insert cDNA sequence contained in pc901HISgalpha12.
XX
KW Human; anorectic; antidiabetic; antihypertensive; hypothalamus;
KW G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
KW clobopobia; anorexia nervosa; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 15..2102
FT /tag= a
FT /product= "pc901HISgalpha12 protein"

XX
PN W02003030936-A1.
XX
PD 17-APR-2003.
XX
PF 02-OCT-2002; 2002MO-JP10250.
XX
PR 02-OCT-2001; 2001JP-00306872.
XX
PA (SUMU) SUMITOMO PHARM CO LTD.
XX
PI Suguru E, Teuchida A, Yamanaka M, Taiji M;
XX
DR WPI: 2003-35486/33.
DR P-PSDB; ABR56305.
XX
PT Inhibitors of expression or activity of G-protein coupled receptor 901
PT for treatment of lifestyle-related diseases and clobopobia.
XX
PS Example 5; Page 75-78; 91pp; Japanese.
XX
CC The present invention relates to novel remedies for the treatment of
CC diseases containing as an active component an inhibitor of the expression
CC or activity of hypothalamus-expressed G-protein coupled receptor 901 and
CC for treatment of clobopobia containing as an active component a
CC potentiator of the expression or activity of G-protein coupled receptor
CC 901. The diseases which can be treated include obesity, diabetes and
CC hyperlipaemia, and clobopobia (anorexia nervosa). The present sequence
CC was used to illustrate the invention
XX
SQ Sequence 2216 BP; 456 A; 688 C; 657 G; 415 T; 0 U; 0 other;

Query Match 99.7%; Score 999; DB 8; Length 2216;
Best Local Similarity 100.0%; Pred. No. 1.6e-193;
Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGAGAGTGAAGCTGTTCACAGGACAGGAGCTGGTGGAGAGCTGTGCTGCGAG 60
DB 15 ATGTGAGAGTGAAGCTGTTCACAGGACAGGAGCTGGTGGAGAGCTGTGCTGCGAG 74
QY 61 GACCTGAGAGTGGAGCTGTTCACAGGAGCTGTGCTGCGAGCTGTGCTGCGAGCTGTG 120
DB 75 GACCTGAGAGTGGAGCTGTTCACAGGAGCTGTGCTGCGAGCTGTGCTGCGAGCTGTG 134
QY 121 GCGCTGTGTTCAACAGCGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 135 GCGCTGTGTTCAACAGCGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 194
QY 181 ATGCCGAGGTGATCTTTGTCACATGGAGGTGGCAAGGCTGTGCTGCTGCTGCTGCTG 240
DB 195 ATGCCGAGGTGATCTTTGTCACATGGAGGTGGCAAGGCTGTGCTGCTGCTGCTGCTG 254
QY 241 CTTGTGACCTGCTCGGCCCCCGAGCTCCCGGTGGCGCTGTGAGTGTGGCGGAA 300
DB 255 CTTGTGACCTGCTCGGCCCCCGAGCTCCCGGTGGCGCTGTGAGTGTGGCGGAA 314
QY 301 GTCCACGTGGCACTGCAAGATCCCTTCAATGTGTCTTCACTGTGGCCATGTATCCACC 360
DB 315 GTCCACGTGGCACTGCAAGATCCCTTCAATGTGTCTTCACTGTGGCCATGTATCCACC 374
QY 361 GCGCTGTGAGCTTCCACCACTACATGAGCGTGGCACTGCGCGGCACTACATGTGGCAGC 420
DB 375 GCGCTGTGAGCTTCCACCACTACATGAGCGTGGCACTGCGCGGCACTACATGTGGCAGC 434
QY 421 GTGTAACAACAGCGGCACTGTGTGGCGCTGTGTGGGTGGCGCGCTGTGACCAAGCTTC 480
DB 435 GTGTAACAACAGCGGCACTGTGTGGCGCTGTGTGGGTGGCGCGCTGTGACCAAGCTTC 494
QY 481 TCTCTGCTGCTTTTCAATCTGAGCCATGTGTCCACCGCGGCTTGAAGTGGCCAG 540
DB 495 TCTCTGCTGCTTTTCAATCTGAGCCATGTGTCCACCGCGGCTTGAAGTGGCCAG 554
QY 541 ATGCAAAACGCAAAAGTGGCGGAGCGCAAGCTGTGTTCATCGGCTTACGTGTGCCAGCA 600

```

Db      555 ATGAGAACCGAGAGCTGCGACGCCACGCTGGTTCATCGGCTACGCTGTCACGCA 614
QY      601 CTGGCCACCTCTACGCGCTGCTGCTACTCTCCCGCTCGGACGGAGACAGCCCTTG 660
Db      615 CTGGCCACCTCTACGCGCTGCTGCTACTCTCCCGCTCGGACGGAGACAGCCCTTG 674
QY      661 GACCGGACACAGGCGCGCTGAGACCTCGGACACAGAGCTGCTGGTGGCCACCGTATGC 720
Db      675 GACCGGACACAGGCGCGCTGAGACCTCGGACACAGAGCTGCTGGTGGCCACCGTATGC 734
QY      721 ACCGAGTTTGGCTCTGAGACGCCACACTATCTGATCTGCTGGGACACAGCTCATCATC 780
Db      735 ACCGAGTTTGGCTCTGAGACGCCACACTATCTGATCTGCTGGGACACAGCTCATCATC 794
QY      781 TCGGAGGAGAAAGCCGCTGAGACGCCACACTACTGCGGGCTACTGCACTTTTGAAAGATTTC 840
Db      795 TCGGAGGAGAAAGCCGCTGAGACGCCACACTACTGCGGGCTACTGCACTTTTGAAAGATTTC 854
QY      841 TCCAACTCTGCGCTTCTCTCAGACAGTTTGTGACACCACTTCTTACCGCTACATGAAC 900
Db      855 TCCAACTCTGCGCTTCTCTCAGACAGTTTGTGACACCACTTCTTACCGCTACATGAAC 914
QY      901 CAGAGCTTCCCAAGCAAGCTTCAACGCGCTGATGAAAAAGCTGCGCTGCGGGACCGGAC 960
Db      915 CAGAGCTTCCCAAGCAAGCTTCAACGCGCTGATGAAAAAGCTGCGCTGCGGGACCGGAC 974
QY      961 TGCTCCCGGACCAACATGGGGGTGACAGAGGTGCTGGCG 999
Db      975 TGCTCCCGGACCAACATGGGGGTGACAGAGGTGCTGGCG 1013

```

RESULT 14

ACCT0859
ID ACCT0859 standard; DNA; 2258 BP.

ACCT0859;

20-NOV-2003 (first entry)

Insert cDNA sequence contained in pc901HISgalphal6.

Human; anorectic; antidiabetic; antilipemic; hypothalamus;
G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
cibophobia; anorexia nervosa; ds.

Unidentified.

Key Location/Qualifiers

FT CDS 15..2159 /tag=a /product="pc901HISgalphal6 protein"

MO2003030936-A1.

17-APR-2003.

02-OCT-2002; 2002WO-JP010250.

02-OCT-2001; 2001JP-00306872.

(SUMU) SUMITOMO PHARM CO LTD.

Sugaru E, Tsuchida A, Yamanaka M, Taiji M,

WPI; 2003-354886/33.

P-PSDB; ABR56304.

Inhibitors of expression or activity of G-protein coupled receptor 901
for treatment of lifestyle-related diseases and cibophobia.

Example 5; Page 69-73; 91pp; Japanese.

CC The present invention relates to novel remedies for the treatment of
CC diseases containing as an active component an inhibitor of the expression
CC or activity of hypothalamus-expressed G-protein coupled receptor 901 and
CC for treatment of cibophobia containing as an active component a
CC potentiator of the expression or activity of G-protein coupled receptor
CC 901. The diseases which can be treated include obesity, diabetes and
CC hyperlipaemia, and cibophobia (anorexia nervosa). The present sequence
XX was used to illustrate the invention

SQ Sequence 2258 BP; 436 A; 728 C; 687 G; 407 T; 0 U; 0 Other;

Query Match 99.7%; Score 999; DB 8; Length 2258;
Best Local Similarity 100.0%; Pred. No. 1,66-193;

Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCTGAGCTGTTCAACGCGACAGGGCTGTGAGAGCTGCTGCTGCCAG 60

Db 15 ATGAGAGCTGAGCTGTTCAACGCGACAGGGCTGTGAGAGCTGCTGCTGCCAG 74

QY 61 GACCTGACGCTGGGGCTGTCATCTGTTGCTGCTGGGCTGTGCTGGGCTGCCAGTG 120

Db 75 GACCTGACGCTGGGGCTGTCATCTGTTGCTGCTGGGCTGTGCTGGGCTGCCAGTG 134

QY 121 GGCCTGTGCTAACAGCCCTGCTGCTGGGCACTTACACAGCAAGCCAGATGACC 180

Db 135 GGCCTGTGCTAACAGCCCTGCTGCTGGGCACTTACACAGCAAGCCAGATGACC 194

QY 181 ATGCGGAGAGTGTACTTTGTCAACATGAGCAGTGTGAGGCTGTGCTAGCGCCCTGAGCC 240

Db 195 ATGCGGAGAGTGTACTTTGTCAACATGAGCAGTGTGAGGCTGTGCTAGCGCCCTGAGCC 254

QY 241 CCTGTGACCTGCTGCGGCCCCCGAGCTCCCGGTGGGGGCTGTGAGTGTGGGCGGCA 300

Db 255 CCTGTGACCTGCTGCGGCCCCCGAGCTCCCGGTGGGGGCTGTGAGTGTGGGCGGCA 314

QY 301 GTCCAGTGTGAGCTGACAGATCCCTTCAATGTGTCTTCACTGATGTGGCCATGACTCCACC 360

Db 315 GTCCAGTGTGAGCTGACAGATCCCTTCAATGTGTCTTCACTGATGTGGCCATGACTCCACC 374

QY 361 GCCCTGTGAGCTGTGACCACTACATGAGGCTGTGACCTGCGGAGCCTTACATGGCCAGC 420

Db 375 GCCCTGTGAGCTGTGACCACTACATGAGGCTGTGACCTGCGGAGCCTTACATGGCCAGC 434

QY 421 GTGTCAACACAGCGGCAAGTGTGGGCTTGTGTGGGGTGGCGGCTGTACACAGCTTC 480

Db 435 GTGTCAACACAGCGGCAAGTGTGGGCTTGTGTGGGGTGGCGGCTGTACACAGCTTC 494

QY 481 TCTTGCTGCTCTTCTACATGTGACGCCATGTGTCAACCCGCGGCTAGAGTGGCCCAAG 540

Db 495 TCTTGCTGCTCTTCTACATGTGACGCCATGTGTCAACCCGCGGCTAGAGTGGCCCAAG 554

QY 541 ATGAGAAAGGAGAAAGCTGCGAGCCAGCCAGCTGTGTTTCACTGCGCTACGTGTGCCAGCA 600

Db 555 ATGAGAAAGGAGAAAGCTGCGAGCCAGCCAGCTGTGTTTCACTGCGCTACGTGTGCCAGCA 614

QY 601 CTGGCCACCTCTACGCGCTGCTGCTACTCTCCCGCTCGGACGAGAGACAGCCCTTG 660

Db 615 CTGGCCACCTCTACGCGCTGCTGCTACTCTCCCGCTCGGACGAGAGACAGCCCTTG 674

QY 661 GACCGGACACAGGCGCGCTGAGACCCCTCGGACACAGGCTGTGTTGCCACCGCTGTGC 720

Db 675 GACCGGACACAGGCGCGCTGAGACCCCTCGGACACAGGCTGTGTTGCCACCGCTGTGC 734

QY 721 ACGAGTTTGGGCTCTGAGGCCACACTATCTGATCTGCTGGGGACACAGGTATCATC 780

Db 735 ACGAGTTTGGGCTCTGAGGCCACACTATCTGATCTGCTGGGGACACAGGTATCATC 794

QY 781 TCGGAGAGGAAAGCCGCTGAGACGCCACACTACTGCGGGCTACTGCACTTTTGAAAGATTTC 840

Db 795 TCGGAGAGGAAAGCCGCTGAGACGCCACACTACTGCGGGCTACTGCACTTTTGAAAGATTTC 854

QY 841 TCCAACTCTGCGCTTCTCAGACAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900

This Page Blank (uspto)


```

Db      369  CTTGTCACCTGCTGCGCCCCCGAGCTCCGGTGGGCGCTGTGAGTGTTGGCGCGAA 428
QY      301  GTCCAGTGGCACTGAGATCCCTTCAATGTCTCTCACTGTGGCCATGATCTCCACC 360
Db      429  GTCCAGTGGCACTGAGATCCCTTCAATGTCTCTCACTGTGGCCATGATCTCCACC 488
QY      361  GCCCTCTGAGCTTCGACCACTACATGAGCGTGA CTGCGCGGACCTACATGAGCCAG 420
Db      489  GCCCTCTGAGCTTCGACCACTACATGAGCGTGA CTGCGCGGACCTACATGAGCCAG 548
QY      421  GTGTACAAACAGCGGCACTGTGCGGCTTGTGTGGGTTGGCGCGCTGTGACCACTTC 480
Db      549  GTGTACAAACAGCGGCACTGTGCGGCTTGTGTGGGTTGGCGCGCTGTGACCACTTC 608
QY      481  TCTGCTGCTCTTCAACATCTGCAACCATGTGTCCACC CGCGCTGTAAGTGGCCAG 540
Db      609  TCTGCTGCTCTTCAACATCTGCAACCATGTGTCCACC CGCGCTGTAAGTGGCCAG 668
QY      541  ATGCAGAAACGAGAACTGCGCGACGCGACGCTGTGTTCTATCGGCTACGTTGGCCAGCA 600
Db      669  ATGCAGAAACGAGAACTGCGCGACGCGACGCTGTGTTCTATCGGCTACGTTGGCCAGCA 728
QY      601  CTGGCCACCTCTTACGCGCTGTGTACTCTTCCGCGTCCGAGGAGCAACGCCCTTG 660
Db      729  CTGGCCACCTCTTACGCGCTGTGTACTCTTCCGCGTCCGAGGAGCAACGCCCTTG 788
QY      661  GACCGGGAACAGCGGCGCGCTGTGAGCGCTGTGAGCAACAGGCTGTGTGGCCACCGGTTC 720
Db      789  GACCGGGAACAGCGGCGCGCTGTGAGCGCTGTGAGCAACAGGCTGTGTGGCCACCGGTTC 848
QY      721  ACGCAATTTGGGCTGTGAGCGCACTATCTGATCTCTGCGGACACAGGTCATCATC 780
Db      849  ACGCAATTTGGGCTGTGAGCGCACTATCTGATCTCTGCGGACACAGGTCATCATC 908
QY      781  TGGCGAGGGAAGCCCTGTGAGCGCACTACTCTGCGGCTACTGCACTTTGTGAAGATTTC 840
Db      909  TGGCGAGGGAAGCCCTGTGAGCGCACTACTCTGCGGCTACTGCACTTTGTGAAGATTTC 968
QY      841  TCCAAACTCTGCGCTTCTCAAGAGCTTGTGACACCACTTCTACCGCTACATGAAAC 900
Db      969  TCCAAACTCTGCGCTTCTCAAGAGCTTGTGACACCACTTCTACCGCTACATGAAAC 1028
QY      901  CAGAGCTTCCCGAGCAAGCTCCACCGCTGATGATAAAAGCTCCCTGCGGAGCCGCGCAC 960
Db      1029  CAGAGCTTCCCGAGCAAGCTCCACCGCTGATGATAAAAGCTCCCTGCGGAGCCGCGCAC 1088
QY      961  TGTCTCCCGGACCACTGGGGGTGACGAGGTGCTGGCGTAG 1002
Db      1089  TGTCTCCCGGACCACTGGGGGTGACGAGGTGCTGGCGTAG 1130

RESULT 2
US-09-409-778-1
; Sequence 1, Application US/09409778
; Patent No. 6472173
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Yeung, George
; TITLE OF INVENTION: A NOVEL CHEMOKINE RECEPTOR OBTAINED FROM
; FILE REFERENCE: 20411-742CON2 (now 28110/36057B)
; CURRENT FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US/09/409,778
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/236,166
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: US 09/106,800
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PasteSeq for Windows Version 3.0
; SEQ ID NO 1

```

```

; LENGTH: 300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-409-778-1
Query Match
Best Local Similarity 99.6%; Pred. No. 9.3e-43;
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      772  GTATCATATCTGCGGAGGGAAGCCGTGAGACGACACTACCTTGTGAGCA 831
Db      1  GTATCATATCTGCGGAGGGAAGCCGTGAGACGACACTACCTTGTGAGCA 60
QY      832  AAGGATTTCTCCAACTCTGCGCTTCTCCAGAGCTTGTGAGCA 891
Db      61  AAGGATTTCTCCAACTCTGCGCTTCTCCAGAGCTTGTGAGCA 120
QY      892  TACATGAACCAAGGCTTCCCGACAGCTCCACGCTGATGATAAAAGCTGCCCTGCGGG 951
Db      121  TACATGAACCAAGGCTTCCCGACAGCTCCACGCTGATGATAAAAGCTGCCCTGCGGG 180
QY      952  GACCGGCACTGTCTCCCGAGACCAATGGGGGTGACAGGTGCTGGCGTAG 1002
Db      181  GACCGGCACTGTCTCCCGAGACCAATGGGGGTGACAGGTGCTGGCGTAG 231

RESULT 3
US-09-170-496D-3
; Sequence 3, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-170-496D-3
Query Match
Best Local Similarity 42.4%; Pred. No. 8.9e-07;
Matches 380; Conservative 0; Mismatches 517; Indels 0; Gaps 0;

QY      52  GCTGGCCAGACCTGCAAGTGGGCTGTGACATGTTGTGCTGGGCTGTGGTGGG 111
Db      22  GGTGTCACAGTGAACGTGGCGGTGACACCACTCTTCCGCTATCTACATCTTGTTC 81
QY      112  GTGCAAGTGGGCTGTGTAACAAGCCCTGCTGGTGTGCACTACACAGCAAGGCC 171
Db      82  ATGCGGCTGGGCTGGCCCACTAAGTGGCTGTGTGGGCGGCTACCGCAAGTGA 141
QY      172  AGCATGACCATGCGGAGCGTGTACTTGTCAACATGAGCAAGTGGAGCGCTGTGACG 231
Db      142  CAGGCAACGAGCTGGGGGTGTACTGATGAACCTGACAGCATGCGGCACTGTGTATC 201
QY      232  GCCCTGGCCCTGTGTGACACTGTGGGCCCCGAGGCTCCCGGTGGGCGGTGTGAAGTGTG 291
Db      202  TGAACGCTGCGCTGTGGGTGTGACTTCTGTGACCAAGCAACTGATTCACGCGCCC 261
QY      292  GCGCGCAAGTCCAGTGGGCACTGACAGATCCCTTCAATGTGTCTCACTGTGGCGCATG 351
Db      262  GGTGCTGCAAGCTTGTGGTTATCTTCAACCAATATCTACATCAAGATGCGCTTC 321
QY      352  TACTCCACCGGCTGTGAGGCTGTGACCACTACATGAGGCTGTGACGCGGAGCTTAC 411

```

Db 322 CTGTGTCGATCTCGGTGACCGGCTACCTGGCTGTGGCCACCACCTCCGCTTCCGCCGC 381
 Qy 412 ATGGCCAGGGTGTAAACAACGCGGCAAGTGTGGGCTGTGGGGTGGCGGCTGTG 471
 Db 382 CTGGCGGGGTCAAGACCGCGTGGCGGTGAGTCCGGTGTGGGCAACGAGCTGGGC 441
 Qy 472 ACCAGCTTCTCTGCTGTCTTCTACATCTGCAAGCAGTGTGCAACCCGCGGCTAGAG 531
 Db 442 GCCAATCGGGCGGCTGTTCATGACAGAGCTTTCGAGACCGGCTACAAACACACTTC 501
 Qy 532 TGGCCCAAGATGCAAGACGAGAGCTGCCAGCCACGCTGTGTTCATCGGCTACGTG 591
 Db 502 TGGCTTGAAGAAAGTCCCAATGGAAGCTGGGTGGCTGAGTGAACCTCTATCGGCTGTC 561
 Qy 592 GTGGCCAGCACTGGCCCAACCCCTTACGGGCTGTACTCTCCCGGCTCCGAGAGAGAC 651
 Db 562 GTGGGCTTCTTCCGCTGGGGGCTCATGCTGTCTGTGACCGGGGCAATCTGGCGGCC 621
 Qy 652 AGCCCTGGACCGGAGACAGCGGCGGCTGGAGCCCTCGGACACAGGCTGTGTGGCC 711
 Db 622 GTGGGGGGGAGCGTGTTCACCGAGCCGCGAGAGAGCCAAAGTCAAGCGGCTGGCCCTC 681
 Qy 712 ACCGTGTGACGAGATTGGGCTGTGAGCGCCACACTATCTGATCTGTGGGCAACAG 771
 Db 682 AGCTCATGCGCATCGTGTGTGTCTGTGGCGCCCTATCAAGTGTCTGTGTCCCGC 741
 Qy 772 GTTCATATCTCGGAGAGGAGCCCGGTGAGACGACACTACCTGGGGCTACGTCACTTGTG 831
 Db 742 AGGCGCATCTACTGTGGCGCGCCCTGGAGCTGGCGCTTCAGAGAGCGCGCTTTTCTGCA 801
 Qy 832 AAGGATTTCTCCAACTCTGAGCTTCTCTCAGCAGCTTTGTGACACACTTCTCTACCG 891
 Db 802 TACCAAGCTCACTGCTTTCACACAGCTTCACTGTGTGGCGAGCCCACTCTCTACTGC 861
 Qy 892 TACATGACCAAGCTTCCCGACCAAGCTCCAAAGCTGTGAAAAAGCTGCCCTGC 948
 Db 862 CTGGTCAAGAGGGCGCGGACGATGTGGCCAGAGGCGCTGCAACACTGTCTCCGC 918
 Db
 RESULT 4
 US-09-016-434-1429
 / Sequence 1429, Application US/09016434
 / Patent No. 6500938
 / GENERAL INFORMATION:
 / APPLICANT: Janice Au-Young
 / APPLICANT: Jeffrey J. Seilhammer
 / TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 / TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 / NUMBER OF SEQUENCES: 1490
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 / STREET: 3174 PORTER DRIVE
 / CITY: PALO ALTO
 / STATE: CALIFORNIA
 / COUNTRY: USA
 / ZIP: 94304
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/016,434
 / FILING DATE: HEREWITTH
 / CLASSIFICATION:
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER:
 / FILING DATE:
 / CLASSIFICATION:
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Zeller, Karen J.
 / REGISTRATION NUMBER: 37,071
 / REFERENCE/DOCKET NUMBER: PA-0002 US

/ TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (650) 855-0555
 / TELEFAX: (650) 845-4166
 / INFORMATION FOR SEQ. ID NO: 1429:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1365 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / IMMEDIATE SOURCE:
 / LIBRARY: GENBANK
 / CLONE: 9598152
 / US-09-016-434-1429
 Query Match 7.0%; Score 69.8; DB 4; Length 1365;
 Best Local Similarity 42.4%; Pred. No. 9.4e-07;
 Matches 380; Conservative 0; Mismatches 517; Indels 0; Gaps 0;
 Qy 52 GCCTGCGAGACCTGCGAGCTGGGCTGTGCACTGTGTGCTGTGGGCTGTGTGGGC 111
 Db 238 GCGTGCAGGTGAGACTGCGCGGTGACACACTCTTCCGCACTCCCTTACATCTTGTG 297
 Qy 112 GTGCCAGTGGGCTGTGTCAACAGCCCTGTGTGTGTGGCCAACTTACAGCAAGGCC 171
 Db 298 ATGCGGTGGGGCTGCCCAACACTGCTGTGTGTGGGCGGCTTACCGCCAGGTGCA 357
 Qy 172 AGATATACCATGCCGAGAGTATCTTTGTCAACATGGAGTGGCAGGCGGTGTGTGAG 231
 Db 358 CAGCGCAAGAGTGGGCGGTCTACATGATGACCTGAGATGCGCACTGCTGTATCATC 417
 Qy 232 GCCCTGGCCCTGTGTGACCTGTGCGGCCGCCAGCTCCCGGTGGGCGGTGTGAGTGTG 291
 Db 418 TGCAGCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 477
 Qy 292 GCGCGGAAGTCCAGTGTGCACTGACATGATCCCTTCAATGTGTCTTCACTGTGTG 351
 Db 478 GGGTCTGCAAGCTCTTGGGTTTCACTTCAACCAATATATCAATCAAGACTGGCTTC 537
 Qy 352 TACTCCAGCGCTGTGTGAGCTGTGACACTATGACATGTGACATGTGCGGAGCTTAC 411
 Db 538 CTGTGTGATCTCGGTGTGAGACCGTACCTGAGCTGTGGCCCAACCACTGCGCTTGC 597
 Qy 412 ATGGCAGGTGTAAACAACGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 471
 Db 598 CTGGCGCGGTCAAGACCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 657
 Qy 472 ACCAGCTTCTCTGCTGTGTCTTCTTCAATGTGAGCATGTGTGCAACCGCGGCTAG 531
 Db 658 GCCAATCGGGCGGCTGTTCATGACAGAGCTTTCGAGACCGCTAACCAACACACTTC 717
 Qy 532 TGGCCCAAGATGCAAGACGAGAGCTGCCAGCCAGCTGTGTGTGTGTGTGTGTGTGT 591
 Db 718 TGGTTGAAGAAAGTCCCAATGGAAGCTGGGTGGCTGGAATGAACCTTATCGGGTTC 777
 Qy 592 GTGGCAGCACTGGCCCAACCCCTTACCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 651
 Db 778 GTGGGCTTCTTCCGCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 837
 Qy 652 AGCCCTGTGAGACCGGAGACGCGGCGGCTGTGAGCCCTCGGACACAGGCTGTGTGT 711
 Db 838 GTGGGGGGGAGGTGTTCACCGAGCCGAGAGAAAGCCAAAGTCAAGGCGGCTGGCCCT 897
 Qy 712 ACCGTGTGACGAGATTGGGCTGTGTGAGCCCACTATGTGTATCTGTGTGTGTGTGT 771
 Db 898 AGCTCATGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 957
 Qy 772 GTTCATATCTCGGAGAGGAGCCCGGTGAGCGCAACTATCTGAGGCTGTGTGTGTGT 831
 Db 958 AGGCGCATCTACTGTGGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017
 Qy 832 AAGGATTTCTCCAACTCTGTGCGCTTCTCAGAGCTTTGTGACACTTCTTCTACCG 891
 Db 1018 TACCAAGCTCACTGCTTTCACACAGCTTCACTGTGTGTGTGTGTGTGTGTGTGTGT 1077

QY 892 TACATGAACGAGGCTTCCCGAGCAAGCTCCACGCGCTGATGAAAAAGCTGCGCTGC 948
Db 1078 CTGGTCAACGAGGGGCGCCCGCAGCGCATGTGGCCAAAGGCCCTGCACACTGCTCCGC 1134

RESULT 5
US-09-170-496D-165
Sequence 165, Application US/09170496D
Patent No. 6555339
GENERAL INFORMATION:
APPLICANT: Benhan, Dominic P.
APPLICANT: Chalmers, Derek T.
APPLICANT: Liaw, Chen W.
TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
FILE REFERENCE: AREN-0040
CURRENT APPLICATION NUMBER: US/09/170,496D
CURRENT FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 294
SOFTWARE: Patentin version 3.1
SEQ ID NO 165
LENGTH: 1089
TYPE: DNA
ORGANISM: Homo sapiens
US-09-170-496D-165

Query Match 6.8%; Score 68.2; DB 4; Length 1089;
Best Local Similarity 42.3%; Pred. No. 2e-06;

Matches 379; Conservative 0; Mismatches 518; Indels 0; Gaps 0;

QY 52 GCTTCGACGAGACCTGCAAGCTGGGGCTGTCACTGTTGCTGCTGGGGCTGGTGGGC 111
Db 22 GGCCTGCACCTGGACTCGCCGCTGGAACCACTCTTCCGCACTCTTACATCTTGTTC 81

QY 112 GTGCGAGTGGGCGCTGTCTACAAAGCCCTGCTGCTGAGCCAACTACACAGCAAGCC 171
Db 82 ATGCGCGTGGGGGCTGCCACCACTGCTGTGGCTGTGGGGGAGCTTACAGCTTGTTC 141

QY 172 AGCATGACATGCGCGAGCTGTACTTGTCAACATGAGCAAGCTGAGCTGTGTCTAGC 221
Db 142 CAGGCAACGAGCTGGGGGCTACTGCTGATGACCTGAGCTGCGGCACTGTGTACTATC 201

QY 232 GCGCTGGCCCTGTGTGACCTGTGTGGCCCCCGAGCTCCGGGTGGGGCTGTGGAAGTGTG 291
Db 202 TGCACTGCTGCGCTGTGGGTGAGTACTTCTGTGACCAACAACTGGAATCAAGGCC 261

QY 292 GCGCGGAGATGCACTGAGCACTGAGATCCCTTCAATGTCTCTCACTGTGTGTGTGT 351
Db 262 GGGTCTGCAAGCTCTTGGGTTCAATTTTCAACCAATATCTACATCAAGATGGCTTTC 321

QY 352 TACTCAACGCGCTGTGTGAGCTGTGACCACTACATGAGGGTGTGACTGCGCGAGCTTAC 411
Db 322 CTGTGTGCAATCTGTGGTGTGACCGCTACTGAGCTGTGCGCCCACTCCCTTTCGCGCG 381

QY 412 ATGCGCAGCGCTGTCAACAGCGCGCAGTGTGGGCTTCTGTGTGGGTGGGGGCTGTGCTG 471
Db 382 CTGGCGCGCTCAACAGCGCGCTGTGGGTGTGAGCTTCCGTGTGTGTGGGCAAGGAGTGGC 441

QY 472 ACCAGCTTCTCTGCTGTCTTCTTACATGTGACATGTGTCACACCGGCGCTGTAGAG 531
Db 442 GCGAATCTGGGCGCCCTGTTCATGAGAGAGCTTTCGAGACGCTTACACACACTTTC 501

QY 532 TGGCCCAAGATGCAAGAAAGCAAGCTGCCAGCCAGCTGTGTGTCTATCGCTAGTGT 591
Db 502 TGCTTTGAGAGTTCCCATGAGAGGCTGTGGGTGTGAGTGTGAACTTATCGGGTGTTC 561

QY 592 GTGCGAGACTGGGCAACCTCTAGCGCGTGTGTACTCTCCGCGTCCGAGGGAGAGAC 651
Db 562 GTGGGCTTCTCTTCTCCGTGGGCGCTATGTGCTGTGTGTACCGGGGCAATCTGTGGGCC 621

QY 652 ACGCCCTTGAACCGGAGACAGGCGCGCTGTGAGCCCTGTGGCAACAAGGCTGTGTGTG 711

Db 622 GTGGGGGCAAGCTGTCTCCAGCGCCAGGAGAAAGGCAAGATCGCGGCTGCGCTTC 681
QY 712 ACCGTGTGACGAGCTTGGGCTCTGAGCGCACTATCTGATCTGCTGGGCGACAG 771
Db 682 AGCCATATCGCCATCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 741

QY 772 GTCATCATCTGCGAGGAGAAAGCGGTGTGACGACACTACTGTGGGCTACTGCACTTGTG 831
Db 742 AGCCCATCTACTGTGGGCGCCCTGTGGGACTGTGGGCTTTCAGAGAGCGCGTCTTTCTGCA 801

QY 832 AAGATTTCTCCAACTCTGTGCTTCTTCCAGCACTTGTGACCACTTCTTACCGC 891
Db 802 TACCACACTACTGTGGCTTTCACCAAGCTTCACTGTGTGGGAGCCCATCTCTACTGC 861

QY 892 TACATGAACGAGGCTTCCCGAGCAAGCTCCACGCGCTGATGAAAAAGCTGCGCTGC 948
Db 862 CTGGTCAACGAGGGGCGCCGAGCGATGTGGCCAAAGGCCCTGCACAACTGCTCCGC 918

RESULT 6

US-09-299-843A-65
Sequence 65, Application US/09299843A
Patent No. 6107475

GENERAL INFORMATION:

APPLICANT: Godieka, Ronald
APPLICANT: Gray, Patrick W.
APPLICANT: Schweikart, Vicki L.
TITLE OF INVENTION: No. 6107475e1 Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: 6100 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA

ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,843A
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/088,337
FILING DATE: 01-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,848
FILING DATE: 17-NOV-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: J111 E. Uhl
REGISTRATION NUMBER: 43,213
REFERENCE/DOCKET NUMBER: 27866/32059B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX:

SEQUENCE CHARACTERISTICS:
LENGTH: 2085 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 177..1310

US-09-299-843A-65

Query Match 6.3%; Score 62.8; DB 3; Length 2085;
 Best Local Similarity 43.7%; Pred. No. 3.9e-05;
 Matches 374; Conservative 0; Mismatches 477; Indels 5; Gaps 2;

QY 118 GTGGGCTGTGCTCAACGCGCTGTGTGCTGCGCAACCTACAGACAGAGGCGCATG 177
 DB 387 GTGGGCTGTGCTGCGCAACGCGCTGTGTGCTGCGCAACCTACAGACAGAGGCGCATG 446
 QY 178 ACCATGCGGACGTGATCTTGTCAACATGAGCAAGGCTGTGTGCTGCGCAACCTG 237
 DB 447 ACCATGACGAGATACCTACCTGCTCAACCTGCGCGTGGACAGACATCTTTCTCTATT 506
 QY 238 GCCCTGTGACCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 297
 DB 507 CTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 562
 QY 298 GAAGTCCA-CGTGCACTGCAAGATCCCTTCAATGTGCTCACTGCTGCGCATGTACTC 356
 DB 563 TAAAGGACATTTTGGCATCTATTAAGTTAAGCTTTTCAAGCGGATGCTGCTCTATG 622
 QY 357 CACCGCCCTGTGAGCTGACCACTACATGACGCTGCACTGCGCGCACTACATGAC 416
 DB 623 CATCAGCATTGACCGCTACGATGACGATGCTGCGCGCGCGCGCGCGCGCGCGCGCG 682
 QY 417 CAGGTGTACACACGCGGCACTGTGCGCGCTGTGTGCGCGCGCGCGCGCGCGCGCG 476
 DB 683 CCGCGTGTCTCTATCAGCAAGCTGTCTGTGTGCGCGCATGTGATGCTGCGCGCTCT 742
 QY 477 CTTCCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 536
 DB 743 CTTCATCTCGGAGCTGCTCTCAACGCGCTCTCAAGACAGCGCGCAAGACAGCTGAG 802
 QY 537 CAAGATGACAGACGCAAGACGCTGCGCAAGCTGCTGTCTCTCTCTCTCTCTCTCTCT 596
 DB 803 ATGCTCACTGATGATGCTGCGCAAGGAGGCTGTGATCACCATCCAGTGGCGCGATGT 862
 QY 597 AGCATGCGCAACCTCTACGCGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 656
 DB 863 TTTTGGTTCCTAGTGTCTATGCTGCTGCTATGATGCTTCTGCTACTCTATATCATCCGTAC 922
 QY 657 CCTGACCGGAGACAGGCGCGCGCTGAGCCCTGCGCAACAGAGTGTGAGTGCACCT 716
 DB 923 CTTCCTCTCAAGCAAGCACTTTTGAAGCGAACAAGGCTATCATATGCGCTGTGT 982
 QY 717 GTGACGCACTTTGAGCTGTGACGCGCACTATCTGATCTGTGTGCGCGCAACGCTCAT 776
 DB 983 GGTATGCTTATATGCTTTCAGAGTGGCTCTCAATGAGGCTGTGTGTCTGACGCTGCG 1042
 QY 777 CATCTGCGAGGAGAGCGCGCTGAGCGCACTACCTGTGCGCTACTGCACTTTGTGAAGA 836
 DB 1043 CAACCTTCACATCAACCAATAGACGCTGGAACAGCAAGACAGCTCAACCTTGTCTATGA 1102
 QY 837 TTTCTCCAACTCTGCGCTTCTCCAGAGCTTTGTGACACACCTTCTCAACGCTCAT 896
 DB 1103 CGTCACTTCAAGCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1162
 QY 897 GAACCAAGGCTTCCCAAGCAAGCTCAACGCGCTGATGAAAAAGCTGCTGCGGAGCG 956
 DB 1163 CGAGCTCAAGTTCCGAGCGACCTCTTCAAGCTCTTCAAGGACTTGGGCTGCTCAAGCA 1222
 QY 957 GCACTGCTCCCGGAC 972
 DB 1223 GGAACGCTCCGGCAC 1238

RESULT 7
 US-09-088-337B-65
 ; Sequence 65; Application US/09088337B
 ; Patent No. 6348574
 ; GENERAL INFORMATION:
 ; APPLICANT: Godtiska, Ronald

Gray, Patrick W.
 Schweikart, Vicki L.
 TITLE OF INVENTION: No. 6348574el Seven Transmembrane Receptors
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/088,337B
 FILING DATE: 01-Jun-1998
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/153,848
 FILING DATE: 17-NOV-1993
 APPLICATION NUMBER: US 07/977,452
 FILING DATE: 17-NOV-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6348574and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 31794
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 65:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2085 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 177..1310
 SEQUENCE DESCRIPTION: SEQ ID NO: 65:
 US-09-088-337B-65
 Query Match 6.3%; Score 62.8; DB 3; Length 2085;
 Best Local Similarity 43.7%; Pred. No. 3.9e-05;
 Matches 374; Conservative 0; Mismatches 477; Indels 5; Gaps 2;

QY 118 GTGGGCTGTGCTCAACGCGCTGTGTGCTGCGCAACCTACAGACAGAGGCGCATG 177
 DB 387 GTGGGCTGTGCTGCGCAACGCGCTGTGTGCTGCGCAACCTACAGACAGAGGCGCATG 446
 QY 178 ACCATGCGGACGTGATCTTGTCAACATGAGCAAGGCTGTGTGCTGCGCAACCTG 237
 DB 447 ACCATGACGAGATACCTACCTGCTCAACCTGCGCGTGGACAGACATCTTTCTCTATT 506
 QY 238 GCCCTGTGACCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 297
 DB 507 CTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 562
 QY 298 GAAGTCCA-CGTGCACTGCAAGATCCCTTCAATGTGCTCACTGCTGCGCATGTACTC 356
 DB 563 TAAAGGACATTTTGGCATCTATTAAGTTAAGCTTTTCAAGCGGATGCTGCTCTATG 622
 QY 357 CACCGCCCTGTGAGCTGACCACTACATGACGCTGCACTGCGCGCACTACATGAC 416
 DB 623 CATCAGCATTGACCGCTACGATGACGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 682
 QY 417 CAGGTGTACACACGCGGCACTGTGCGCGCTGTGTGCGCGCGCGCGCGCGCGCGCGCG 476

Oy	159	CAcAGcAAGcCAGcATGAcCATGcCcGAGcGTGAcCTTTGcCAAcATGAGcAGTGGcCAGG	218
Db	153	GGGcATcAAcCTTcCGGcCTcGAcCTTcAcCGGGcGGATcGcTGAATcGAcGTcGAcGAAcCA	212
Oy	219	CCTGcGTcTcAcGcGcCCTcTGGcCCCTcGTGcAcCTcGTcCGAcCCcCCGAcCTcCCGcGTGGcC	278
Db	213	AcCCGcCCAcTcCTcGGcCAAGcGTGcGTGcTcCAATcGTcGGcCCcCGGcCTAcGAGcAGcCCGT	272
Oy	279	GcTGTGcAGcTGTGGcCGcGAGcATcCAcGTGcAcCTGcAGATcCCcCTTcCAATGTGTcCTc	338
Db	273	GcTGcAAGcCTTcCGGcCAcGcCCcCGcAGcCTcGTGTGGcATcGcTcCAcGAGcAGAcCCGGA	332
Oy	339	AcTGTGTGGcCAcTGTAcCTcCAcCGcCCTcGTGcAGcCTcGAcCAcTAcATcGAcGcGTGAcCT	398
Db	333	AcTGcGAGcAAAGcTGTGcCTAcCGcCTTcCAcGAcGAGcCGAcTGGcCAAcCCGcGcGAAcCT	392
Oy	399	GcCCcCGAcCTAcATGcGCcAcGTGTGTAcAACAcCGcGGAcAGcTGTGTGGcCTTcGTGTGGGG	458
Db	393	GAACcCGGcTcAGTAcGTcCGcCCcCGcAGcGTGGcTGAAGAcTGTGGcCAcAGGGcGGcCTT	452
Oy	459	TGGcCGcGTcGTGAcCAGcCTTcCTcCGcCTcCTcTTAcATcTGTcCAGcCAcTGTGTcCAC	518
Db	453	CGGcATGTcCTcTGGcGcCTcGGGGcGGcATcTcTGTAcGTcCGGcCTTcCGcCTTcCAcGTGAA	512
Oy	519	CCGcGGcGTAcGTGcCGcCAAGcATGcCAGAcCGAcAGAcCTcCCAcGcCAcAGcCTGTGT	578
Db	513	GTGTGGcCCTcGTGTGcATcCTcTGTGTGTGTcCAGcATcCGAcTAcGTGTGcATGTGGcGTGT	572
Oy	579	CATcGGcCTAcGTGTGcCCAGcAcCTGGcCAcCCcCTTAcCGcGTGGcTGT	626
Db	573	GTcGTcCTTcCAcGTGTAcCTTcGAcCTAcCGcTGTcCTcTGGcCCcGTGTGT	620

```

RESULT 13
US-09-252-991A-6054/c
Sequence 6054, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIORITY APPLICATION NUMBER: US 60/074,788
PRIORITY FILING DATE: 1998-02-18
PRIORITY APPLICATION NUMBER: US 60/094,190
PRIORITY FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 6054
LENGTH: 984
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6054

Query Match      5.7%; Score 57.6; DB 4; Length 984;
Best Local Similarity 44.3%; Pred. No. 0.00048;
Matches 234; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

99 CCTGGTGTGGGCGGTGCGCAAGTGGGGCTGTGCTACACAGCCCTGCTGTGTGGCCAACT 158
Db CGTGGCGGTGGCCGTCACCCCTGATCTGACCCGATCCCTCGGGCAAGCTGTTCACCAA 851

159 ACACAGCAAGGCGACGATGACCATGCGCGGAGGTGATCTTTGTCAACATGGCAGTGGCAGG 218
Db GGGCATCACTTCGGCCCTCGAATTACCGGCGGTACGCTGATTCAGCTGACGTACGAACA 791

219 CTTGGTGTCAAGCGCCCTTGTCGACCTGCTCGGCCCCCGAGACTCCGGTGGGC 278
Db ACCGGCCGATCTCGGCAAGTGCGGTGATCAATGTCGGCCCGGCTACGAGACCGCGT 731

279 GCTGTGAGTGTGGGCGGCGAAGTCCACGTGGCACTGCAGATCCCTTCAATGTGTCTCT 338

```

Db 730 GGTGACGACTTTCGGGAGAGCCCGGAGAGTCTCGTGTGGATTGGCCAGCGAGACCAGGA 6713

Qy 339 ACTGTGGCCATGTACTCCATCCGCCCTCTGTAGCTTGACCACTAATCGAGCTGCAT 3398

Db 670 ACTGGGCAAGAGGTGTGCTTACCGCTTCGACAGGCCATGCCGAGCAACCCGGGAACT 6111

Qy 399 GCCGGGAACTCATGGCAGCGTGTACAAACCGCGGACGTCGTGGGACTTCGTGTGGG 4588

Db 610 GAAGCCGGGTGAGTACGTGAGCCCGCAGGTGGGTGAAAGAACTGGCGCAACCGAGGGGCT 5511

Qy 459 TGGCGGCTGTGACCACTTCTCTCGCTCTTTTAACTGACGCAATGTTCAC 518

Db 550 CGGCAATGCTTCGGCGCTGGGCGGAGATCTCTGTACGTCGGCTTCGCTTCAGTGGAA 4911

Qy 519 CCGCGGGCTAGATGCGCCCAAGATGACAGAGAGTCCGACCGCACGCTGGTGT 578

Db 490 GTTGTGCTTCGGTGCAATCTCTCGCTGTGCAACATGCAATCAATCGTAGTGGCGTCT 4111

Qy 579 CATCGGCTACGTGTGTCAGCACTGAGCAACCTCTACGCGCTGGTGTCT 626

Db 430 GTGCTTCTTCAGGTACCTTGAACATCAACGTCCTCGCCGCTGTGCT 383

RESULT 14
 US-08-540-650B-6
 Sequence 6, Application US/08540650B
 Patent No. 639325
 GENERAL INFORMATION:
 APPLICANT: HINDMA, Shuji
 APPLICANT: FUJII, Ryo
 APPLICANT: FUKUSUMI, Shoji
 APPLICANT: OHTAKI, Tetsuya
 APPLICANT: HOSOYA, Masaki
 APPLICANT: OHGI, Kazuhiko
 APPLICANT: ONGA, Haruo
 TITLE OF INVENTION: GALANIN RECEPTOR PROTEIN, PRODUCTION AND USE THEREOF
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
 STREET: 130 WATER STREET
 CITY: BOSTON
 STATE: MA
 COUNTRY: US
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/540,650B
 FILING DATE: 11-OCT-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 7-134412
 FILING DATE: 31-MAY-1995
 APPLICATION NUMBER: 6-326610
 FILING DATE: 28-DEC-1994
 APPLICATION NUMBER: 6-247599
 FILING DATE: 13-OCT-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: RESNICK, DAVID S
 REGISTRATION NUMBER: 34,235
 REFERENCE/DOCKET NUMBER: 45901
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-523-3400
 TELEFAX: 617-523-6440
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1047
 TYPE: Nucleic acid
 STRANDEDNESS: Double

Search completed: December 15, 2004, 02:56:16
Job time : 87.2892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:25:49 ; Search time 511.891 Seconds
(without alignments)
10792.424 Million cell updates/sec

Title: US-09-995-225B-1

Perfect score: 1002

Sequence: 1 atctggcgccgcagcgcgtc.....tgcagcgcgtcgtcgcgtag 1002

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1002	100.0	1002	9	US-09-995-225-1
2	1002	100.0	1002	10	US-09-990-940-3
3	1002	100.0	1002	10	US-09-995-225-1
4	1002	100.0	1402	15	US-10-292-798-1015
5	1002	100.0	1561	15	US-10-017-161-1201
6	1002	100.0	1614	15	US-10-101-148-2
7	1002	100.0	1778	10	US-09-791-932-54
8	1002	100.0	1785	15	US-10-225-567a-679
9	1002	100.0	1928	14	US-10-165-844-14
10	1000.4	99.8	1750	10	US-09-791-932-31
11	1000.4	99.8	1878	9	US-09-895-686-9
12	1000.4	99.8	1893	16	US-10-296-115-435

13	577	57.6	2651	9	US-09-764-847-1782	Sequence 1782, Ap
14	577	57.6	2651	14	US-10-092-154-1782	Sequence 1782, Ap
15	358.6	35.8	368	10	US-09-918-995-8117	Sequence 8117, Ap
16	350.6	35.0	461	10	US-09-918-995-30890	Sequence 30890, A
17	304	30.3	497	9	US-09-895-686-59	Sequence 59, Appl
18	275	27.4	225	9	US-09-895-686-32	Sequence 32, Appl
19	236.4	23.6	238	9	US-09-895-686-30	Sequence 30, Appl
20	196	19.6	196	15	US-10-101-148-1	Sequence 1, Appl
21	178	17.8	930	9	US-09-764-847-1781	Sequence 1781, Ap
22	178	17.8	930	14	US-10-092-154-1781	Sequence 1781, Ap
23	169	16.9	294	9	US-09-895-686-61	Sequence 61, Appl
24	166	16.6	259	9	US-09-895-686-31	Sequence 31, Appl
25	162.6	16.2	262	9	US-09-895-686-60	Sequence 60, Appl
26	134	13.4	255	9	US-09-895-686-27	Sequence 27, Appl
27	122	12.2	281	9	US-09-895-686-29	Sequence 29, Appl
28	89.4	8.9	363	9	US-09-895-686-28	Sequence 28, Appl
29	69.8	7.0	1089	15	US-10-251-385-3	Sequence 3, Appl
30	69.8	7.0	1089	15	US-10-267-811-1	Sequence 17, Appl
31	69.8	7.0	1365	15	US-10-101-510-17	Sequence 1429, Ap
32	69.8	7.0	1365	16	US-10-305-720-1429	Sequence 113, App
33	69.8	7.0	1365	17	US-10-283-975A-113	Sequence 1, Appl
34	69.8	7.0	2693	10	US-09-850-948-1	Sequence 1, Appl
35	69.8	7.0	2693	15	US-10-273-575-1	Sequence 1, Appl
36	69.8	7.0	2696	15	US-10-225-567a-272	Sequence 272, App
37	69.8	7.0	2980	15	US-10-101-510-456	Sequence 456, App
38	68.2	6.8	1089	15	US-10-251-385-165	Sequence 165, App
39	62.8	6.3	1137	15	US-10-004-113-21	Sequence 21, Appl
40	62.8	6.3	2072	9	US-09-815-937-19	Sequence 19, Appl
41	62.8	6.3	2072	15	US-10-004-113-20	Sequence 20, Appl
42	60.8	6.1	1068	14	US-10-237-563-48	Sequence 48, Appl
43	58.4	5.8	1170	17	US-10-225-567a-537	Sequence 537, App
44	58.4	5.8	1170	17	US-10-775-920-14	Sequence 14, Appl
45	58.4	5.8	1215	15	US-10-220-883-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-995-225-1

Sequence 1, Application US/09995225

Publication No. US20020193584A1

GENERAL INFORMATION:

APPLICANT: Chen, Huoping

APPLICANT: Chu, Zhi Liang

APPLICANT: Dang, Huang T.

APPLICANT: Lowitz, Kevin P.

TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human

TITLE OF INVENTION: Receptors

FILE REFERENCE: AREN-0308

CURRENT APPLICATION NUMBER: US/09/995, 225

CURRENT FILING DATE: 2001-11-26

PRIOR APPLICATION NUMBER: 09/170,496

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: PCT/US99/23938

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: 60/253,404

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/255,366

PRIOR FILING DATE: 2000-12-12

PRIOR APPLICATION NUMBER: 60/270,286

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,365

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/270,266

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,032

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,358

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,356

PRIOR FILING DATE: 2001-04-06

```
/ PRIOR APPLICATION NUMBER: 60/290,917
/ PRIOR FILING DATE: 2001-05-14
/ PRIOR APPLICATION NUMBER: 60/309,208
/ PRIOR FILING DATE: 2001-07-31
/ NUMBER OF SEQ ID NOS: 67
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO: 1
/ LENGTH: 1002
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: No. US20020193584A1e1 Sequence
US-09-995-225-1
```

```
Query Match      100.0%; Score 1002; DB 9; Length 1002;
Best Local Similarity 100.0%; Pred. No. 7e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGTGAGCTGCAGCTGTTCAACGGGCAAGGGCTGTGGAGAGCTGCTGCTGCGCAG 60
Db      1 ATGTGAGCTGCAGCTGTTCAACGGGCAAGGGCTGTGGAGAGCTGCTGCTGCGCAG 60

QY      61 GACCTGAGCTGGGGGCTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db      61 GACCTGAGCTGGGGGCTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

QY      121 GGCTGTGCTACAAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db      121 GGCTGTGCTACAAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

QY      181 ATGCGGAGCTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db      181 ATGCGGAGCTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

QY      241 CCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db      241 CCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

QY      301 GTCCAGTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db      301 GTCCAGTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360

QY      361 GCGCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db      361 GCGCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY      421 GTGTACAACAGCGGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db      421 GTGTACAACAGCGGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY      481 TCCTGCTGCTGCTTCTACATCTGCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db      481 TCCTGCTGCTGCTTCTACATCTGCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

QY      541 ATGCGAAGAGCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db      541 ATGCGAAGAGCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY      601 CTGGCAACCTCTACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db      601 CTGGCAACCTCTACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

QY      661 GACCGGAGCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db      661 GACCGGAGCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

QY      721 AGCGAGTTTGGGCTCTGCAAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      721 AGCGAGTTTGGGCTCTGCAAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY      781 TGGCGAGGGAGCGCGTGGAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db      781 TGGCGAGGGAGCGCGTGGAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
```

```
QY      841 TCCAAACTCTGCGCTTCTCCAGAGAGCTTTGTGACACCACTTCTTACCGGCTACATGAAAC 900
Db      841 TCCAAACTCTGCGCTTCTCCAGAGAGCTTTGTGACACCACTTCTTACCGGCTACATGAAAC 900

QY      901 CAGAGCTTCCCGAGCAAGCTTCCAAAGGCTGATGATAAAGAGCTGCGGGAGACCGGGAC 960
Db      901 CAGAGCTTCCCGAGCAAGCTTCCAAAGGCTGATGATAAAGAGCTGCGGGAGACCGGGAC 960

QY      961 TGTCTCCCGAGACCAATGGGGGTGACAGAGGTGCTGCGCTAG 1002
Db      961 TGTCTCCCGAGACCAATGGGGGTGACAGAGGTGCTGCGCTAG 1002
```

RESULT 2

```
US-09-990-940-3
/ Sequence 3, Application US/0990940
/ Publication No. US20030027252A1
/ GENERAL INFORMATION:
/ APPLICANT: Tian, Hui
/ APPLICANT: Zhao, Jiaqiang
/ APPLICANT: Chen, Jin-Long
/ APPLICANT: Cutler, Gene
/ APPLICANT: An, Songzhu
/ APPLICANT: Dai, Kang
/ APPLICANT: Gupta, Jamila S.
/ APPLICANT: Tularik Inc.
/ TITLE OF INVENTION: No. US20030027252A1e1 Receptors
/ FILE REFERENCE: 018781-007410US
/ CURRENT APPLICATION NUMBER: US/09/990,940
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: US 60/252,841
/ PRIOR FILING DATE: 2000-11-22
/ PRIOR APPLICATION NUMBER: US 60/257,636
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 60/261,377
/ PRIOR FILING DATE: 2001-01-12
/ PRIOR APPLICATION NUMBER: US 60/279,554
/ PRIOR FILING DATE: 2001-03-28
/ PRIOR APPLICATION NUMBER: US 60/280,696
/ PRIOR FILING DATE: 2001-03-29
/ SOFTWARE: Patent Ver. 2.1
/ SEQ ID NO: 3
/ LENGTH: 1002
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1002)
/ OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR339
US-09-990-940-3
```

```
Query Match      100.0%; Score 1002; DB 10; Length 1002;
Best Local Similarity 100.0%; Pred. No. 7e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ATGTGAGCTGCAGCTGTTCAACGGGCAAGGGCTGTGGAGAGCTGCTGCTGCGCAG 60
Db      1 ATGTGAGCTGCAGCTGTTCAACGGGCAAGGGCTGTGGAGAGCTGCTGCTGCGCAG 60

QY      61 GACCTGAGCTGGGGGCTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db      61 GACCTGAGCTGGGGGCTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

QY      121 GGCTGTGCTACAAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db      121 GGCTGTGCTACAAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

QY      181 ATGCGGAGCTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db      181 ATGCGGAGCTGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
```

QY	241	CTGTGACCTGCTCGGCCCCCGAGCTCCGGTGGGCGCTGTGGAGTGTGGGCGGCGAA	300
Db	241	CTGTGACCTGCTCGGCCCCCGAGCTCCGGTGGGCGCTGTGGAGTGTGGGCGGCGAA	300
QY	301	GTTCACGTGGCAGCTGCAGATCCCTTCAATGTGTCTCACTGTGTGGCATTACTAC	360
Db	301	GTTCACGTGGCAGCTGCAGATCCCTTCAATGTGTCTCACTGTGTGGCATTACTAC	360
QY	361	GCCCTGTGAGCTCGACCACTACATGAGCGTGCACTTCCGCGGACCTTACATGCGCAGC	420
Db	361	GCCCTGTGAGCTCGACCACTACATGAGCGTGCACTTCCGCGGACCTTACATGCGCAGC	420
QY	421	GTGTACAAACCGGGGCACTGTGTGGGCTTCTGTGTGGGCTGTGGGCGGCTGTGACCACTTC	480
Db	421	GTGTACAAACCGGGGCACTGTGTGGGCTTCTGTGTGGGCTGTGGGCGGCTGTGACCACTTC	480
QY	481	TCTGTGTGTCTTCTACATCTGAGCATGTGTCCACCCTGGCGCTAGAGTGCAGCAAG	540
Db	481	TCTGTGTGTCTTCTACATCTGAGCATGTGTCCACCCTGGCGCTAGAGTGCAGCAAG	540
QY	541	ATGCAAGACGACAGAGCTGCCGACGCACTGTGTCTTCTGTGTGTGTGTGTGTGTGTGT	600
Db	541	ATGCAAGACGACAGAGCTGCCGACGCACTGTGTCTTCTGTGTGTGTGTGTGTGTGTGT	600
QY	601	CTGGGCACTCTTACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	660
Db	601	CTGGGCACTCTTACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	660
QY	661	GACCGGGAACCGGGCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	720
Db	661	GACCGGGAACCGGGCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	720
QY	721	ACGCACTTTGGGCTCTGTGACGCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGT	780
Db	721	ACGCACTTTGGGCTCTGTGACGCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGT	780
QY	781	TGCGAGAGGAAGCCGTGTGACGCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGT	840
Db	781	TGCGAGAGGAAGCCGTGTGACGCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGT	840
QY	841	TCCAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	900
Db	841	TCCAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	900
QY	901	CAGAGCTTTCCCGACAGCTCCAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	960
Db	901	CAGAGCTTTCCCGACAGCTCCAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	960
QY	961	TGCTTCCCGGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1002
Db	961	TGCTTCCCGGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1002
RESULT 3			
US-09-995-225-1			
; Sequence 1, Application US/0995225			
; Publication No. US20030139588A9			
; GENERAL INFORMATION:			
; APPLICANT: Chen, Ruoping			
; APPLICANT: Chu, Zhi Liang			
; APPLICANT: Dang, Huong T.			
; APPLICANT: Lowitz, Kevin P.			
; APPLICANT: Pidge, Cameron			
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G			
; TITLE OF INVENTION: Receptors			
; FILE REFERENCE: AREN-0308			
; CURRENT APPLICATION NUMBER: US/09/995,225			
; CURRENT FILING DATE: 2001-11-26			
; PRIOR APPLICATION NUMBER: 09/170,496			
; PRIOR FILING DATE: 1998-10-13			
; PRIOR APPLICATION NUMBER: PCT/US99/23938			
; PRIOR FILING DATE: 1998-10-13			
; PRIOR APPLICATION NUMBER: 60/253,404			

```

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/299,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;
; LENGTH: 1002
;
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139589a9e1 Sequence
US-09-995-225-1

```

Query Match	100.0%;	Score 1002;	DB 10;	Length 1002;
Best Local Similarity	100.0%;	Pred. No. 7e-262;		
Matches 1002;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

QY 601 CTGGCCACCCCTCTACGCGCTGATGCTACTCTCCCGCTCCGACAGGAGGACACGCCCCCTG 660
Db 601 CTGGCCACCCCTCTACGCGCTGATGCTACTCTCCCGCTCCGACAGGAGGACACGCCCCCTG 660
QY 661 GACCGGAGACACGCGCCGCTGAGACCCCTCGGACACAGGCTGTGTGTGACCCGCTGTC 720
Db 661 GACCGGAGACACGCGCCGCTGAGACCCCTCGGACACAGGCTGTGTGTGACCCGCTGTC 720
QY 721 ACCGAGTTGGGCTCTGGAAGGCGCACTATCTGATCTCTGTGGGGGACACAGGTATCATC 780
Db 721 ACCGAGTTGGGCTCTGGAAGGCGCACTATCTGATCTCTGTGGGGGACACAGGTATCATC 780
QY 781 TCGCAGGAGGAGCCGCTGAGACGACACTACCTGCGGCTACTGCACTTTGTGAAGATTTC 840
Db 781 TCGCAGGAGGAGCCGCTGAGACGACACTACCTGCGGCTACTGCACTTTGTGAAGATTTC 840
QY 841 TCCAACTCTGCGCTTCTGCAAGAGTTTGTGACACCACTTTCTACCGCTACATGAAC 900
Db 841 TCCAACTCTGCGCTTCTGCAAGAGTTTGTGACACCACTTTCTACCGCTACATGAAC 900
QY 901 CAGAGCTTCCCGACGACGCTCAACCGGCTGATGAAAAAGCTGCGCGGGGACCGGAC 960
Db 901 CAGAGCTTCCCGACGACGCTCAACCGGCTGATGAAAAAGCTGCGCGGGGACCGGAC 960
QY 961 TGCTCCCGGACCACTAGGGGGTGCAGAGGTGCTGCGGTAG 1002
Db 961 TGCTCCCGGACCACTAGGGGGTGCAGAGGTGCTGCGGTAG 1002

RESULT 4
US-10-292-798-1015
Sequence 1015, Application US/10292798
Publication No. US2003023583A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: KATYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1015
LENGTH: 1402
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE: (1) .. (1402)
LOCATION: (1) .. (1402)
NAME/KEY: CDS
LOCATION: (201) .. (1202)
US-10-292-798-1015

Query Match 100.0%; Score 1002; DB 15; Length 1402;
Best Local Similarity 100.0%; Pred. No. 7,3e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGGAGCTGAGCTGTTTCAACGGACAGAGGCTGTGTGAGAGACTGCTTCCCTGCCAG 60
Db 201 ATGTGGAGCTGAGCTGTTTCAACGGACAGAGGCTGTGTGAGAGACTGCTTCCCTGCCAG 260
QY 61 GACCTGACAGCTGCGGCTGTCTACTGTGTGCTGTGCGGCTGGTGGTGGAGTGCACAGT 120
Db 261 GACCTGACAGCTGCGGCTGTCTACTGTGTGCTGTGCGGCTGGTGGTGGAGTGCACAGT 320

QY 121 GGCCTGTGCTAACAGCCCTGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 321 GGCCTGTGCTAACAGCCCTGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380
QY 181 ATGCGGAGGATGATCTTTGTCAACATGAGCAGTGGAGGCTGGTGTGCTGAGCCCTGGCC 240
Db 381 ATGCGGAGGATGATCTTTGTCAACATGAGCAGTGGAGGCTGGTGTGCTGAGCCCTGGCC 440
QY 241 CCTGTGACCTGTGCTGCGCCCTGCGAGCTTCCGCTGGTGGAGTGTGGGCGGCA 300
Db 441 CCTGTGACCTGTGCTGCGCCCTGCGAGCTTCCGCTGGTGGAGTGTGGGCGGCA 500
QY 301 GTCCAGTGGGACCTGACATGCCCTTCAATGTGTCTCATGTGTGCTGATGCTACCTCACC 360
Db 501 GTCCAGTGGGACCTGACATGCCCTTCAATGTGTCTCATGTGTGCTGATGCTACCTCACC 560
QY 361 GCCCTGTGAGCTGACACCTACATGACAGCTGTGACTCCGCGAGCTTACATGAGCCAC 420
Db 561 GCCCTGTGAGCTGACACCTACATGACAGCTGTGACTCCGCGAGCTTACATGAGCCAC 620
QY 421 GTGTACACACCGGACAGTGTGCGGCTTCTGTGTGGGTGTGCGCTGTGACCACTTC 480
Db 621 GTGTACACACCGGACAGTGTGCGGCTTCTGTGTGGGTGTGCGCTGTGACCACTTC 680
QY 481 TCTGCTGCTGCTTCTTACATGACAGCTGATGTGACCGCGCTGAGTGGGCGCAG 540
Db 681 TCTGCTGCTGCTTCTTACATGACAGCTGATGTGACCGCGCTGAGTGGGCGCAG 740
QY 541 ATGCAAAACGACAGAGCTGCGGACGCGACGCTGTGTATCTGCTGCTGCTGCTGCTGCTG 600
Db 741 ATGCAAAACGACAGAGCTGCGGACGCGACGCTGTGTATCTGCTGCTGCTGCTGCTGCTG 800
QY 601 CTGGCCACCCCTCTACGCGCTGTGTGCTTCTCCGCTGCGAGGAGGACACGCCCCCTG 660
Db 801 CTGGCCACCCCTCTACGCGCTGTGTGCTTCTCCGCTGCGAGGAGGACACGCCCCCTG 860
QY 661 GACCGGAGACGCGGCGCTGAGAGCCCTGCGACACAGAGCTGTGTGAGCCACCGGTGTC 720
Db 861 GACCGGAGACGCGGCGCTGAGAGCCCTGCGACACAGAGCTGTGTGAGCCACCGGTGTC 920
QY 721 ACGAGTTTGGGCTCTGAGCGCCACACTATGTGATCTGTGTGGGACACAGGTATCATC 780
Db 921 ACGAGTTTGGGCTCTGAGCGCCACACTATGTGATCTGTGTGGGACACAGGTATCATC 980
QY 781 TCGCAGGAGGAGCCGCTGAGACGCACTACCTGTGGGCTACTGCACTTTGTGAAGATTTC 840
Db 981 TCGCAGGAGGAGCCGCTGAGACGCACTACCTGTGGGCTACTGCACTTTGTGAAGATTTC 1040
QY 841 TCCAACTCTGCGCTTCTGCAAGAGTTTGTGACACCACTTCTTACCGCTACATGAAC 900
Db 1041 TCCAACTCTGCGCTTCTGCAAGAGTTTGTGACACCACTTCTTACCGCTACATGAAC 1100
QY 901 CAGAGCTTCCCGACGACGCTCAACGGCTGATGAAAAAGCTGCGGCTGCGGGACCGGAC 960
Db 1101 CAGAGCTTCCCGACGACGCTCAACGGCTGATGAAAAAGCTGCGGCTGCGGGACCGGAC 1160
QY 961 TGCTCCCGGACCACTAGGGGGTGCAGAGGTGCTGCGGTAG 1002
Db 1161 TGCTCCCGGACCACTAGGGGGTGCAGAGGTGCTGCGGTAG 1202

RESULT 5
US-10-017-161-1201
Sequence 1201, Application US/10017161
Publication No. US2003014368A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: KATYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161

```

; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1201
; LENGTH: 1561
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1561)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1361)
US-10-017-161-1201

Query Match      100.0%; Score 1002; DB 15; Length 1561;
Best Local Similarity 100.0%; Pred. No. 7, 4e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGTGAGCTGCACTGGTTCACACGGCAGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 60
DB      360  ATGTGAGCTGCACTGGTTCACACGGCAGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 419
QY      61  GACCTGCACTGGGCGCTGTCACTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB      420  GACCTGCACTGGGCGCTGTCACTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 479
QY      121  GGCCTGTGCTACACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB      480  GGCCTGTGCTACACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
QY      181  ATGCCGAGCTGTACTTTGTCAACATGCACTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB      540  ATGCCGAGCTGTACTTTGTCAACATGCACTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 599
QY      241  CCTGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB      600  CCTGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
QY      301  GTCAGAGTGCACATGAGATCCCTTCAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB      660  GTCAGAGTGCACATGAGATCCCTTCAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 719
QY      361  GGCCTGTGAGCTGCACTGCACTGATGAGAGCTGCACTGAGCTGCACTGATGAGAGCTGCA 420
DB      720  GGCCTGTGAGCTGCACTGCACTGATGAGAGCTGCACTGAGCTGCACTGATGAGAGCTGCA 779
QY      421  GTGTACAAACAGCGGCGCAGTGTGCGGCTTGTGTGAGGCTGCGCTGCTGCTGCTGCTGCTG 480
DB      780  GTGTACAAACAGCGGCGCAGTGTGCGGCTTGTGTGAGGCTGCGCTGCTGCTGCTGCTGCTG 839
QY      481  TCTGTGCTGCTTCTTCACTGCTGAGCAGTGTGCAACCGCGCGCTGAGTGTGAGTGTGAG 540
DB      840  TCTGTGCTGCTTCTTCACTGCTGAGCAGTGTGCAACCGCGCGCTGAGTGTGAGTGTGAG 899
QY      541  ATCAGAGACAGAGCTGCGCAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB      900  ATCAGAGACAGAGCTGCGCAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
QY      601  CTGGCCACCTCTTCACTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 660
DB      960  CTGGCCACCTCTTCACTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1019
QY      661  GACCGGGAACAGCGGCGCGCTGAGCGCTGCGACACAGGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB      1020  GACCGGGAACAGCGGCGCGCTGAGCGCTGCGACACAGGCTGCTGCTGCTGCTGCTGCTGCTG 1079
QY      721  ACCGAGTTTGGGCTGTGAGCGCACAATATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB      1080  ACCGAGTTTGGGCTGTGAGCGCACAATATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1139
```

```

QY      781  TCGCAGGAGAGCCCGCTGAGCGCACTACTGTGGGCTACTGCACTTTGTGAGAGATTTC 840
DB      1140  TCGCAGGAGAGCCCGCTGAGCGCACTACTGTGGGCTACTGCACTTTGTGAGAGATTTC 1199
QY      841  TCCAAATCTGCGCTTCTTCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAAC 900
DB      1200  TCCAAATCTGCGCTTCTTCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAAC 1259
QY      901  CAGAGCTTCCCGCAGCACTTCCCAACCGCTGATGATAAAGCTGCTGCGGGAACCGGCAAC 960
DB      1260  CAGAGCTTCCCGCAGCACTTCCCAACCGCTGATGATAAAGCTGCTGCGGGAACCGGCAAC 1319
QY      961  TGCTCCCGGAGACCATGAGGAGGTGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1002
DB      1320  TGCTCCCGGAGACCATGAGGAGGTGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1361

RESULT 6
US-10-101-148-2
; Sequence 2, Application US/10101148
; Publication No. US20030187198A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Yeung, George
; TITLE OF INVENTION: A No. US20030187198A1el Chemokine Receptor Obtained From a cDNA
; TITLE OF INVENTION: Spleen
; FILE REFERENCE: 28110/38317
; CURRENT FILING DATE: US/10/101,148
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-148-2

Query Match      100.0%; Score 1002; DB 15; Length 1614;
Best Local Similarity 100.0%; Pred. No. 7, 4e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGTGAGCTGCACTGGTTCACACGGCAGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 60
DB      129  ATGTGAGCTGCACTGGTTCACACGGCAGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 188
QY      61  GACCTGCACTGGGCGCTGTCACTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB      189  GACCTGCACTGGGCGCTGTCACTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 248
QY      121  GGCCTGTGCTACACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB      249  GGCCTGTGCTACACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 308
QY      181  ATGCCGAGCTGTACTTTGTCAACATGAGAGTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB      309  ATGCCGAGCTGTACTTTGTCAACATGAGAGTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 368
QY      241  CCTGTGCACTGCTGCGCCCGCCGAGCTCCCGGTGAGCGCTGTGAGTGTGAGTGTGAG 300
DB      369  CCTGTGCACTGCTGCGCCCGCCGAGCTCCCGGTGAGCGCTGTGAGTGTGAGTGTGAG 428
QY      301  GTTCAAGTGGCACTGCAAGATCCCTTCAATGTGTCTTCACTGCTGCTGCTGCTGCTGCTGCTG 360
DB      429  GTTCAAGTGGCACTGCAAGATCCCTTCAATGTGTCTTCACTGCTGCTGCTGCTGCTGCTGCTG 488
QY      361  GGCCTGTGAGCTTCAACCACTGATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB      489  GGCCTGTGAGCTTCAACCACTGATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 548
QY      421  GTGTACAAACAGCGGCGCAGTGTGCGGCTTGTGTGAGGCTGCGCTGCTGCTGCTGCTGCTG 480
DB      549  GTGTACAAACAGCGGCGCAGTGTGCGGCTTGTGTGAGGCTGCGCTGCTGCTGCTGCTGCTG 608
```


QY 481 TCCTGCTGCTCTTCTACATCTGACGCCATGTGTCCACCGCGCTAGAGTCCCAAG 540
Db 609 TCCTGCTGCTCTTCTACATCTGACGCCATGTGTCCACCGCGCTAGAGTCCCAAG 668
QY 541 ATGCGAAGGCGAAGAGCTGCGGACGCGACGCTGTGTTCATGCGTACGTGGTGCACAGA 600
Db 669 ATGCGAAGGCGAAGAGCTGCGGACGCGACGCTGTGTTCATGCGTACGTGGTGCACAGA 728
QY 601 CTGCGCACCTCTTACGCGGTGTGTACTCTCCCGGTCGCCGAGGAGAGACACGCCCTTG 660
Db 729 CTGCGCACCTCTTACGCGGTGTGTACTCTCCCGGTCGCCGAGGAGAGACACGCCCTTG 788
QY 661 GACCGGGACACGCGGCGCGCTGGAGCCCTCGGACACACAGCTGTGTGTGCGCACCGGTGTC 720
Db 789 GACCGGGACACGCGGCGCGCTGGAGCCCTCGGACACACAGCTGTGTGTGCGCACCGGTGTC 848
QY 721 ACGCAGTTTGGGCTCTGAGCGCCACACTATCTGATCTCTGCTGGGGACACAGGTATCATC 780
Db 849 ACGCAGTTTGGGCTCTGAGCGCCACACTATCTGATCTCTGCTGGGGACACAGGTATCATC 908
QY 781 TCGCGAGGAAAGCCGCTGAGACGACACTACCTGGGCTACTGCACTTTGTGAAGATTC 840
Db 909 TCGCGAGGAAAGCCGCTGAGACGACACTACCTGGGCTACTGCACTTTGTGAAGATTC 968
QY 841 TCGAAGCTCTGCGCTTCTTCCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
Db 969 TCGAAGCTCTGCGCTTCTTCCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 1028
QY 901 CAGAGCTTCCCGACAGCTTCCAAAGCTGTATGAAAAAGTGCCTCGCGGGACCGGAC 960
Db 1029 CAGAGCTTCCCGACAGCTTCCAAAGCTGTATGAAAAAGTGCCTCGCGGGACCGGAC 1088
QY 961 TGTCTCCCGGACCACTAGGCGGTGACAGAGTGTGCGGTAG 1002
Db 1089 TGTCTCCCGGACCACTAGGCGGTGACAGAGTGTGCGGTAG 1130

RESULT 7
US-09-791-932-54
Sequence 54, Application US/09791932
Publication No. US20030003451A1
GENERAL INFORMATION:
APPLICANT: Vogel, Gabriel
APPLICANT: Parodi, Luis A.
APPLICANT: Hiesch, Ronald R.
APPLICANT: Lind, Peter
APPLICANT: Kayes, Paul S.
APPLICANT: Ruff, Valerie
APPLICANT: Huff, Rita M.
APPLICANT: Wood, Linda S.
TITLE OF INVENTION: No. US20030003451A1 G Protein-Coupled Receptors Cross-Referen
FILE REFERENCE: 00325, US1
CURRENT APPLICATION NUMBER: US/09/791, 932
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/184,305
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,304
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,303
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,397
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,247
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/188,880
PRIOR FILING DATE: 2000-03-13
PRIOR APPLICATION NUMBER: 60/217,369
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/217,370
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/218,492
PRIOR FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: 60/186,810

QY PRIOR FILING DATE: 2000-03-03
Db PRIOR APPLICATION NUMBER: 60/188,064
QY PRIOR FILING DATE: 2000-03-09
Db PRIOR APPLICATION NUMBER: 60/186,457
QY PRIOR FILING DATE: 2000-03-02
Db PRIOR APPLICATION NUMBER: 60/213,861
QY PRIOR FILING DATE: 2000-06-23
Db PRIOR APPLICATION NUMBER: 60/194,344
QY PRIOR FILING DATE: 2000-04-03
Db PRIOR APPLICATION NUMBER: 60/218,337
QY NUMBER OF SEQ ID NOS: 184
Db SOFTWARE: Patent version 3.0
QY SEQ ID NO 54
Db LENGTH: 1778
QY TYPE: DNA
Db ORGANISM: Homo sapiens
QY US-09-791-932-54

Query Match 100.0%; Score 1002; DB 10; Length 1778;
Best Local Similarity 100.0%; Pred. No. 7.5e-282;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 60
Db 19 ATGTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 78
QY 61 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 120
Db 79 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 138
QY 121 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 180
Db 139 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 198
QY 181 ATGTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 240
Db 199 ATGTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 258
QY 241 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 300
Db 259 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 318
QY 301 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 360
Db 319 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 378
QY 361 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 420
Db 379 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 438
QY 421 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 480
Db 439 GACCTGAGCTGTGAGCTGTGTTCAAGCGACACAGGCTGTGTGAGAGCTGCTGCTGCGAG 498
QY 481 TCCTGCTGCTCTTCTACATCTGACGCCATGTGTCCACCGCGCTAGAGTCCCAAG 540
Db 499 TCCTGCTGCTCTTCTACATCTGACGCCATGTGTCCACCGCGCTAGAGTCCCAAG 558
QY 541 ATGCGAAGGCGAAGAGCTGCGGACGCGACGCTGTGTTCATGCGTACGTGGTGCACAGA 600
Db 559 ATGCGAAGGCGAAGAGCTGCGGACGCGACGCTGTGTTCATGCGTACGTGGTGCACAGA 618
QY 601 CTGCGCACCTCTTACGCGGTGTGTACTCTCCCGGTCGCCGAGGAGAGACACGCCCTTG 660
Db 619 CTGCGCACCTCTTACGCGGTGTGTACTCTCCCGGTCGCCGAGGAGAGACACGCCCTTG 678
QY 661 GACCGGGACACGCGGCGCGCTGGAGCCCTCGGACACACAGCTGTGTGTGCGCACCGGTGTC 720
Db 679 GACCGGGACACGCGGCGCGCTGGAGCCCTCGGACACACAGCTGTGTGTGCGCACCGGTGTC 738
QY 721 ACGCAGTTTGGGCTCTGAGCGCCACACTATCTGATCTCTGCTGGGGACACAGGTATCATC 780

Db 739 ACGAGTTTGGGCTGTGACGCCACACTATCTGATCTCTGCGGGGACACAGCTCATCATC 798
Qy 781 TCGCGAGGAAAGCCCGTGGACGACACTACTTGGGGTACTGCACTTTGGAAAGATTTC 840
Db 799 TCGCGAGGAAAGCCCGTGGACGACACTACTTGGGGTACTGCACTTTGGAAAGATTTC 858
Qy 841 TCGAACTCTGCGCTTCTCTCAGAGCTTGTGACACACTTCTCTACGCGTACATGAAC 900
Db 859 TCGAACTCTGCGCTTCTCTCAGAGCTTGTGACACACTTCTCTACGCGTACATGAAC 918
Qy 901 CAGAGCTTCCCGACGAGCTCCACAGCGCTGATGAAAAAGCTGCGCTGCGGGAGACCGGAC 960
Db 919 CAGAGCTTCCCGACGAGCTCCACAGCGCTGATGAAAAAGCTGCGCTGCGGGAGACCGGAC 978
Qy 961 TGTCTCCCGGACCACTAGGGGGTGTGACAGAGTGTGCGCTAG 1002
Db 979 TGTCTCCCGGACCACTAGGGGGTGTGACAGAGTGTGCGCTAG 1020

RESULT 8
US-10-225-567A-679
Sequence 679, Application US/10225567A
Publication No. US20030113798A1
GENERAL INFORMATION:
APPLICANT: LifeSpan Biosciences
APPLICANT: Brown, Joseph P.
APPLICANT: Burner, Glenn C.
APPLICANT: Roush, Christine L.
TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 1920-4-4
CURRENT APPLICATION NUMBER: US/10/225,567A
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/257,144
NUMBER OF SEQ ID NOS: 2292
SOFTWARE: PatentIn version 3.1
SEQ ID NO 679
LENGTH: 1785
TYPE: DNA
ORGANISM: Homo sapiens
US-10-225-567A-679

Query Match 100.0%; Score 1002; DB 15; Length 1785;
Best Local Similarity 100.0%; Pred. No. 7.5e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGAGCTGTGACGCTGTTCAACGACAGAGCTGTGAGAGAGTGTGCTGCTGCGCAG 60
Db 21 ATGTGAGCTGTGACGCTGTTCAACGACAGAGCTGTGAGAGAGTGTGCTGCTGCGCAG 80
Qy 61 GACCTGACGCTGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 81 GACCTGACGCTGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140
Qy 121 GGGCTGTGCTCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 141 GGGCTGTGCTCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 200
Qy 181 ATGCCGGAAGTGTATTTGTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 201 ATGCCGGAAGTGTATTTGTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 260
Qy 241 CCTGTGACCTGTGCTGCGCGCGCGAGAGTCCCGGTGGGCGCTGTGAGTGTGGCGGGGAA 300
Db 261 CCTGTGACCTGTGCTGCGCGCGCGAGAGTCCCGGTGGGCGCTGTGAGTGTGGCGGGGAA 320
Qy 301 GTCGAGTGGCACTGAGATCCGCTTCAATGTGTCTCACTGATGGGCACTGATCTCCACC 360
Db 321 GTCGAGTGGCACTGAGATCCGCTTCAATGTGTCTCACTGATGGGCACTGATCTCCACC 380
Qy 361 GCCCTGTGAGCTGTGACCACTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 381 GCCCTGTGAGCTGTGACCACTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 440

Qy 421 GTGTACAAACAGCGGAGACGCTGCGGCTTGTGTGTGGGGTGTGGGCGCTGACAGCTTC 480
Db 441 GTGTACAAACAGCGGAGACGCTGCGGCTTGTGTGTGGGGTGTGGGCGCTGACAGCTTC 500
Qy 481 TCCTGCTGCTCTTCTACATCTGACAGCAATGTGTCCACCGCGCTAGAGTGCAGAG 540
Db 501 TCCTGCTGCTCTTCTACATCTGACAGCAATGTGTCCACCGCGCTAGAGTGCAGAG 560
Qy 541 ATGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 561 ATGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 620
Qy 601 CTGGCCACCTCTACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 621 CTGGCCACCTCTACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 680
Qy 661 GACCGGGAACAGCGGAGCGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 681 GACCGGGAACAGCGGAGCGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 740
Qy 721 ACGCACTTGGGCTGTGACGCGCACTATCTGATCTGCTGCGGGAACAGGTCATCATC 780
Db 741 ACGCACTTGGGCTGTGACGCGCACTATCTGATCTGCTGCGGGAACAGGTCATCATC 800
Qy 781 TCGCGAGGAAAGCCCGTGGACGACACTACTGATCTGCTGCGGGAACAGGTCATCATC 840
Db 801 TCGCGAGGAAAGCCCGTGGACGACACTACTGATCTGCTGCGGGAACAGGTCATCATC 860
Qy 841 TCGAACTCTGCGCTTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 861 TCGAACTCTGCGCTTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 920
Qy 901 CAGAGCTTCCCGACGAGCTCCACAGCGCTGATGAAAAAGCTGCGCTGCGGGAGACCGGAC 960
Db 921 CAGAGCTTCCCGACGAGCTCCACAGCGCTGATGAAAAAGCTGCGCTGCGGGAGACCGGAC 980
Qy 961 TGTCTCCCGGACCACTAGGGGGTGTGACAGAGTGTGCGCTAG 1002
Db 981 TGTCTCCCGGACCACTAGGGGGTGTGACAGAGTGTGCGCTAG 1022

RESULT 9
US-10-165-844-14
Sequence 14, Application US/10165844
Publication No. US20030017539A1
GENERAL INFORMATION:
APPLICANT: Hodge, Martin R.
APPLICANT: Hunter, John J.
APPLICANT: Rudolph-Owen, Laura
APPLICANT: Welch, Nadine S.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Encoding
FILE REFERENCE: 35800/248302
CURRENT APPLICATION NUMBER: US/10/165,844
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: US 09/088,857
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: US 09/324,465
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: US 09/464,685
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: US 09/741,783
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: US 09/145,745
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: US 09/383,745
PRIOR FILING DATE: 1999-08-26
PRIOR APPLICATION NUMBER: US 09/234,923
PRIOR FILING DATE: 1999-01-21
PRIOR APPLICATION NUMBER: US 09/340,880

```
/ PRIOR FILING DATE: 1998-06-28
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14
/ LENGTH: 1928
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (160)...(1158)
US-10-165-844-14

Query Match      100.0%; Score 1002; DB 14; Length 1928;
Best Local Similarity 100.0%; Pred. No. 7,6e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGTGAGCTGAGCTGTTTCAACGGCAAGGGCTGTGAGAGAGTGCCTGCTCCGCGC 60
DB      160 ATGTGAGCTGAGCTGTTTCAACGGCAAGGGCTGTGAGAGAGTGCCTGCTCCGCGC 219

QY      61 GACCTGAGCTGGGGCTGTACCTGTTGTGTGCTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 120
DB      220 GACCTGAGCTGGGGCTGTACCTGTTGTGTGCTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 279

QY      121 GGCTGTGCTAACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB      280 GGCTGTGCTAACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 339

QY      181 ATGCGGAGCTGTACTTGTTCACATGCGAGTGGCGAGGCTGTGTGTGTGTGTGTGTGTGTGT 240
DB      340 ATGCGGAGCTGTACTTGTTCACATGCGAGTGGCGAGGCTGTGTGTGTGTGTGTGTGTGTGT 399

QY      241 CCTGTGACCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB      400 CCTGTGACCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 459

QY      301 GTCCAGTGGGCACTGCAAGATCCCTTCAATGTGTCTTCACTGTGTGTGTGTGTGTGTGTGTGT 360
DB      460 GTCCAGTGGGCACTGCAAGATCCCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519

QY      460 GTCCAGTGGGCACTGCAAGATCCCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519
DB      520 GTCCAGTGGGCACTGCAAGATCCCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579

QY      520 GCGCTGTGAGCTGTGCAAGATCCCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579
DB      580 GCGCTGTGAGCTGTGCAAGATCCCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 639

QY      421 GTGTACAAACAGCGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB      580 GTGTACAAACAGCGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 639

QY      481 TCCTGCTGTCTTCTTAATCTGAGGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
DB      640 TCCTGCTGTCTTCTTAATCTGAGGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 699

QY      640 TCCTGCTGTCTTCTTAATCTGAGGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 699
DB      700 ATGCAAGACGAGAGAGCTGCGAGCGCGAGCGCGAGCTGTGTGTGTGTGTGTGTGTGTGT 759

QY      541 ATGCAAGACGAGAGAGCTGCGAGCGCGAGCGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 600
DB      700 ATGCAAGACGAGAGAGCTGCGAGCGCGAGCGCGAGCTGTGTGTGTGTGTGTGTGTGTGT 759

QY      700 ATGCAAGACGAGAGAGCTGCGAGCGCGAGCGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 759
DB      760 CTGGCCACCTCTTACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 819

QY      601 CTGGCCACCTCTTACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
DB      760 CTGGCCACCTCTTACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 819

QY      760 CTGGCCACCTCTTACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 819
DB      820 GACCGGGAGACGGGGCGGGCTGTGAGCGCTGTGAGCGAGCGAGCTGTGTGTGTGTGTGTGT 879

QY      820 GACCGGGAGACGGGGCGGGCTGTGAGCGCGCTGTGAGCGAGCGAGCTGTGTGTGTGTGTGT 879
DB      880 AGCAGATTGGGCTGTGAGCGCGCACTATCTGATCTGTGAGGCGAGCGAGCTGTGTGTGTGTGT 939

QY      880 AGCAGATTGGGCTGTGAGCGCGCACTATCTGATCTGTGAGGCGAGCGAGCTGTGTGTGTGTGT 939
DB      940 TGGCAGAGGAGAGCGCGCTGTGAGCGCACTACCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGT 999

QY      940 TGGCAGAGGAGAGCGCGCTGTGAGCGCACTACCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGT 999
DB      841 TCCAAACTCTGCGCTTCTCCAGCAGCTTTGTGACACCACTTCTCTACCGCTACATGAAC 900
```

```
DB      1000 TCCAAACTCTGCGCTTCTCCAGCAGCTTTGTGACACCACTTCTCTACCGCTACATGAAC 1059
QY      901 CAGAGCTTCCCGAGCAAGCTCCCAACGGCTGTATGAAAAAGCTGCTGCGGGAGCCGGCAC 960
DB      1060 CAGAGCTTCCCGAGCAAGCTCCCAACGGCTGTATGAAAAAGCTGCTGCGGGAGCCGGCAC 1119

QY      961 TGTCTCCCGAGCAAGCTTCCCGAGCAAGCTTCCCGAGCAAGCTTCCCGAGCAAGCTTCCCGAG 1002
DB      1120 TGTCTCCCGAGCAAGCTTCCCGAGCAAGCTTCCCGAGCAAGCTTCCCGAGCAAGCTTCCCGAG 1161

RESULT 10
US-09-791-932-31
/ Sequence 31, Application US/09791932
/ Publication No. US20030003451A1
/ GENERAL INFORMATION:
/ APPLICANT: Vogel, Gabriel
/ APPLICANT: Parodi, Luis A.
/ APPLICANT: Hiebsch, Ronald R.
/ APPLICANT: Lind, Peter
/ APPLICANT: Kaytee, Paul S.
/ APPLICANT: Ruff, Valerie
/ APPLICANT: Huff, Rita M.
/ APPLICANT: Wood, Linda S.
/ TITLE OF INVENTION: No. US20030003451A1 G Protein-Coupled Receptors Cross-Referen
/ FID REFERENCE: 00325 US1
/ CURRENT APPLICATION NUMBER: US/09/791,932
/ PRIOR FILING DATE: 2001-02-23
/ PRIOR APPLICATION NUMBER: 60/184,305
/ PRIOR FILING DATE: 2000-02-23
/ PRIOR APPLICATION NUMBER: 60/184,304
/ PRIOR FILING DATE: 2000-02-23
/ PRIOR APPLICATION NUMBER: 60/184,303
/ PRIOR FILING DATE: 2000-02-23
/ PRIOR APPLICATION NUMBER: 60/184,397
/ PRIOR FILING DATE: 2000-02-23
/ PRIOR APPLICATION NUMBER: 60/184,247
/ PRIOR FILING DATE: 2000-02-23
/ PRIOR APPLICATION NUMBER: 60/188,880
/ PRIOR FILING DATE: 2000-03-13
/ PRIOR APPLICATION NUMBER: 60/217,369
/ PRIOR FILING DATE: 2000-07-11
/ PRIOR APPLICATION NUMBER: 60/217,370
/ PRIOR FILING DATE: 2000-07-11
/ PRIOR APPLICATION NUMBER: 60/218,492
/ PRIOR FILING DATE: 2000-07-20
/ PRIOR APPLICATION NUMBER: 60/186,810
/ PRIOR FILING DATE: 2000-03-03
/ PRIOR APPLICATION NUMBER: 60/188,064
/ PRIOR FILING DATE: 2000-03-09
/ PRIOR APPLICATION NUMBER: 60/186,457
/ PRIOR FILING DATE: 2000-03-02
/ PRIOR APPLICATION NUMBER: 60/213,861
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/194,344
/ PRIOR FILING DATE: 2000-04-03
/ PRIOR APPLICATION NUMBER: 60/218,337
/ PRIOR FILING DATE: 2000-07-14
/ NUMBER OF SEQ ID NOS: 184
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 31
/ LENGTH: 1750
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-791-932-31

Query Match      99.8%; Score 1000.4; DB 10; Length 1750;
Best Local Similarity 99.9%; Pred. No. 2e-261; 1; Indels 0; Gaps 0;
Matches 1001; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGTGAGCTGAGCTGTTTCAACGGCAAGGGCTGTGAGAGAGTGCCTGCTCCGCGC 60
```

Db	139	ATGTGAGCTGCAGCTGGATTGTAACAGGCAACAGGAGCTGTGTGGAAGAGACTGCTGACGAC	198
QY	61	GACCTGCAGCTGGGGGCTGTCACTGTTTGTCTGCTGGGCTGTGTGGGCTGTGCAGTG	120
Db	199	GACCTGCAGCTGGGGGCTGTCACTGTTTGTCTGCTGGGCTGTGTGGGCTGTGCAGTG	258
QY	121	GGCCGTGTCTCAAGGCCCTGTGTGGTGTGGGCACTTAACAAGAAAGGCAAGCAATGAAC	180
Db	259	GGCCGTGTCTCAAGGCCCTGTGTGGTGTGGGCACTTAACAAGAAAGGCAAGCAATGAAC	318
QY	181	ATGCCGGAAGTGTACTTTGTCAACATGAGCAGTGGAGAGGCTGTGATCTCAGCCGCTGGCC	240
Db	319	ATGCCGGAAGTGTACTTTGTCAACATGAGCAGTGGAGAGGCTGTGATCTCAGCCGCTGGCC	378
QY	241	CCTGTGCACCTGTCTGGCCCCCGAGACTTCCGGTGGGGCTGTGAGTGTGGGCGGCA	300
Db	379	CCTGTGCACCTGTCTGGCCCCCGAGACTTCCGGTGGGGCTGTGAGTGTGGGCGGCA	438
QY	301	GTCCACGATGGCACTGCAGATCCCTTCAATGTGTCTCACTGTGTGGCAATGAATCTCCAC	360
Db	439	GTCCACGATGGCACTGCAGATCCCTTCAATGTGTCTCACTGTGTGGCAATGAATCTCCAC	498
QY	361	GGCCGTGTGAGGCTGTGACCACTAATGAGCGTGTGACCTGCAGCACTTAATGAGCAGC	420
Db	499	GGCCGTGTGAGGCTGTGACCACTAATGAGCGTGTGACCTGCAGCACTTAATGAGCAGC	558
QY	421	GTGTACAAACAGCGGCAAGTGTGGGGCTTCTGTGTGGGGTGTGGCGCTGTGCACAGCTTC	480
Db	559	GTGTACAAACAGCGGCAAGTGTGGGGCTTCTGTGTGGGGTGTGGCGCTGTGCACAGCTTC	618
QY	481	TCCTGTGTGTCTTTTAACTATCTGCAGCCATGTGTCAACCGGCGGCTGTAGATGGCCCAAG	540
Db	619	TCCTGTGTGTCTTTTAACTATCTGCAGCCATGTGTCAACCGGCGGCTGTAGATGGCCCAAG	678
QY	541	ATGCAGAAACGAGAGCTGCAGCGCAAGCGTGTGTTCATATGAGCTACGTGTGCAGAA	600
Db	679	ATGCAGAAACGAGAGCTGCAGCGCAAGCGTGTGTTCATATGAGCTACGTGTGCAGAA	738
QY	601	CTGGCCACCTCTTACAGCGCTGGTGTACTTCTCCGCGTTCGAGAGGAGGACACGCGCCCTG	660
Db	739	CTGGCCACCTCTTACAGCGCTGGTGTACTTCTCCGCGTTCGAGAGGAGGACACGCGCCCTG	798
QY	661	GACCGGGACACGGGCGGCTGTGAGGCTCTGGGCAACAAGGCTGTGTGTGGCCACCGTGTGC	720
Db	799	GACCGGGACACGGGCGGCTGTGAGGCTCTGGGCAACAAGGCTGTGTGTGGCCACCGTGTGC	858
QY	721	ACGCAAGTTTGGGCTGTGAGCGCAACTATCTGAATCTGTGTGGGCAACAAGGCTGTGTGC	780
Db	859	ACGCAAGTTTGGGCTGTGAGCGCAACTATCTGAATCTGTGTGGGCAACAAGGCTGTGTGC	918
QY	781	TGCGAGAGGGAAGCCCGTGGAGCAACAATACTGTGGGCTTAATGCACTTTGTGAAGATTTC	840
Db	919	TGCGAGAGGGAAGCCCGTGGAGCAACAATACTGTGGGCTTAATGCACTTTGTGAAGATTTC	978
QY	841	TCCAAACTCTGTGGCTTTCTCCAGCAAGCTTTGTGAACAACATTTCTTACCGCTAATGAAC	900
Db	979	TCCAAACTCTGTGGCTTTCTCCAGCAAGCTTTGTGAACAACATTTCTTACCGCTAATGAAC	1038
QY	901	CAGAGCTTCCCGAGCAAGCTTCCAAAGGCGTGAAGAAAGAGCGCCCTGCGGGGAGACCGGCAAC	960
Db	1039	CAGAGCTTCCCGAGCAAGCTTCCAAAGGCGTGAAGAAAGAGCGCCCTGCGGGGAGACCGGCAAC	1098
QY	961	TGCTTCCCGGACCAATGAGGGGTGAGAGAGGTGTGCGGTAG	1002
Db	1099	TGCTTCCCGGACCAATGAGGGGTGAGAGAGGTGTGCGGTAG	1140

DB	Seq ID	Sequence	Score	DB	Length	Matches	Similarity	Pred. No.	Indels	Gaps
DB	1	ATGTGAGACTCAGCTGGTGTCAACGGGCAACGGGCTGTGAGAGAGCTGCTGCGCAG	99.8%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	139	ATGTGAGACTCAGCTGGTGTCAACGGGCAACGGGCTGTGAGAGAGCTGCTGCGCAG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	61	GACCTGACAGCTGGGGCTGTCACTGTGTGCTGCTGGGCTGTGTGTGGGCGTGCAGTG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	199	GACCTGACAGCTGGGGCTGTCACTGTGTGCTGCTGGGCTGTGTGTGGGCGTGCAGTG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	121	GCCCTGTGCTCAACGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	259	GCCCTGTGCTCAACGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	181	ATGCGGAGCTGTACTTGTGTCAACATGAGCAGTGGGAGTGGGAGTGGGAGTGGGAGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	319	ATGCGGAGCTGTACTTGTGTGTCAACATGAGCAGTGGGAGTGGGAGTGGGAGTGGGAGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	241	CCTGTGACAGCTGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	379	CCTGTGACAGCTGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	301	GTCCACGTGGGAGCTGAGATATCCCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	439	GTCCACGTGGGAGCTGAGATATCCCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	361	GCCCTGTGAGCTGTGAGCACTTACATGAGGAGTGGAGTGGGAGTGGGAGTGGGAGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	499	GCCCTGTGAGCTGTGAGCACTTACATGAGGAGTGGAGTGGGAGTGGGAGTGGGAGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	421	GTGTACAAACAGCGGAGAGTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	559	GTGTACAAACAGCGGAGAGTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	481	TCTGTGCTGTCTTCTTACATCTGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	619	TCTGTGCTGTCTTCTTACATCTGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	541	ATGAGAGAGAGAGAGAGTGGGAGAGGAGAGGAGAGTGGGAGAGTGGGAGAGTGGGAG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	679	ATGAGAGAGAGAGAGTGGGAGAGGAGAGGAGAGTGGGAGAGTGGGAGAGTGGGAG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	601	CTGGGAGAGAGAGAGTGGGAGAGGAGAGGAGAGTGGGAGAGTGGGAGAGTGGGAG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	739	CTGGGAGAGAGAGTGGGAGAGGAGAGGAGAGTGGGAGAGTGGGAGAGTGGGAG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	661	GACCGGAGAGAGAGTGGGAGAGGAGAGGAGAGTGGGAGAGTGGGAGAGTGGGAG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	799	GACCGGAGAGAGAGTGGGAGAGGAGAGGAGAGTGGGAGAGTGGGAGAGTGGGAG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	721	ACGAGAGTGGGAGAGTGGGAGAGGAGAGGAGAGTGGGAGAGTGGGAGAGTGGGAG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0
DB	859	ACGAGAGTGGGAGAGTGGGAGAGGAGAGGAGAGTGGGAGAGTGGGAGAGTGGGAG	99.9%	DB 9	1878	0	99.9%	2.1e-261	0	0

QY 781 TCGGAGGAGGAGCCCGGAGCGACACTACCTGCGGCTTCTGCACTTTTGAAGATTTC 840
 Db 919 TCGGAGGAGGAGCCCGGAGCGACACTACCTGCGGCTTCTGCACTTTTGAAGATTTC 978
 QY 841 TCGAACTCTGAGCTTCTCAGCAGCTTTGAGACGACACTTCTTACCGCTTACATGAC 900
 Db 979 TCGAACTCTGAGCTTCTCAGCAGCTTTGAGACGACACTTCTTACCGCTTACATGAC 1038
 QY 901 CAGAGCTTCCCGAGGAGCTCCAGCGCTGATGAAAAGCTGCGCGGGAGCCGAGAC 960
 Db 1039 CAGAGCTTCCCGAGGAGCTCCAGCGCTGATGAAAAGCTGCGCGGGAGCCGAGAC 1098
 QY 961 TGTCTCCCGGAGCGACATGCGGGGTGCGAGGCTGCGGGTAG 1002
 Db 1099 TGTCTCCCGGAGCGACATGCGGGGTGCGAGGCTGCGGGTAG 1140

RESULT 12

US-10-296-115-435
 ; Sequence 435, Application US/10296115
 ; Publication No. US20040053248A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq Inc
 ; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 784PCT
 ; CURRENT APPLICATION NUMBER: US/10/296,115
 ; CURRENT FILING DATE: 2002-11-18
 ; PRIOR APPLICATION NUMBER: US09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; PRIOR APPLICATION NUMBER: US09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 1478
 ; SEQ ID NO 435
 ; LENGTH: 1893
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-296-115-435

Query Match 99.8%; Score 1000.4; DB 16; Length 1893;
 Best Local Similarity 99.9%; Pred. No. 2.1e-261;
 Matches 1001; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTGAGAGCTGCACTGTTGTCACAGGACAGAGGCTGTGAGAGAGCTGCTGCTGCGAG 60
 Db 138 ATGTGAGAGCTGCACTGTTGTCACAGGACAGAGGCTGTGAGAGAGCTGCTGCTGCGAG 197
 QY 61 GACCTGCACTGCGGCTGTCTACTGTTGTGCTGCTGAGGCTGTGAGGAGGCTGCGAGT 120
 Db 198 GACCTGCACTGCGGCTGTCTACTGTTGTGCTGCTGAGGCTGTGAGGAGGCTGCGAGT 257
 QY 121 GGCCTGTGCTAACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 258 GGCCTGTGCTAACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 317
 QY 181 ATGCCGAGAGTGTACTTTGTCAACATGAGCAGTGGAGGCTGTGCTGCTGCTGCTGCTG 240
 Db 318 ATGCCGAGAGTGTACTTTGTCAACATGAGCAGTGGAGGCTGTGCTGCTGCTGCTGCTG 377
 QY 241 CCTGTGCACTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Db 378 CCTGTGCACTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437
 QY 301 GTTCAAGTGTGAGCCTGAGATCCCTTCAATGATGCTTCACTGAGGAGGCTGCTGCTGCTG 360
 Db 438 GTTCAAGTGTGAGCCTGAGATCCCTTCAATGATGCTTCACTGAGGAGGCTGCTGCTGCTG 497
 QY 361 GGCCTGTGAGCCTTGTGAGCCTTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 Db 498 GGCCTGTGAGCCTTGTGAGCCTTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 557
 QY 421 GTTGAACAACGCGGAGCTGTGCGGCTTGTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 480

Db 558 GTTGAACAACGCGGAGCTGTGCGGCTTGTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 617
 QY 481 TGTCTGCTGCTGCTTCTTCAATGTCAGACCAATGTCACACCGCGGCTGAGAGGCGCAG 540
 Db 618 TGTCTGCTGCTGCTTCTTCAATGTCAGACCAATGTCACACCGCGGCTGAGAGGCGCAG 677
 QY 541 ATGCAAGAGCAGAACTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 Db 678 ATGCAAGAGCAGAACTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
 QY 601 CTGAGCACTCTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db 738 CTGAGCACTCTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 797
 QY 661 GACCGGAGACAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Db 798 GACCGGAGACAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857
 QY 721 ACGCAGTTTGGGCTGTGAGACGCACTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Db 858 ACGCAGTTTGGGCTGTGAGACGCACTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 917
 QY 781 TCGCGAGGAGAGCCCTGCTGAGACGCACTTACCTGCGGCTTCTGCTGCTGCTGCTGCTGCTG 840
 Db 918 TCGCGAGGAGAGCCCTGCTGATGCACTTACCTGCGGCTTCTGCTGCTGCTGCTGCTGCTG 977
 QY 841 TCGAACTCTGAGCTTCTCAGACGCTTGTGACACACTTCTTCACTGCTTCACTGCTTCACTG 900
 Db 978 TCGAACTCTGAGCTTCTCAGACGCTTGTGACACACTTCTTCACTGCTTCACTGCTTCACTG 1037
 QY 901 CAGAGCTTCCCGAGGAGCTTCAAGGCTGATGAAAAGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Db 1038 CAGAGCTTCCCGAGGAGCTTCAAGGCTGATGAAAAGCTGCTGCTGCTGCTGCTGCTGCTG 1097
 QY 961 TGTCTCCCGGAGCGACATGCGGGGTGCGAGGCTGCGGGTAG 1002
 Db 1098 TGTCTCCCGGAGCGACATGCGGGGTGCGAGGCTGCGGGTAG 1139

RESULT 13

US-09-764-847-1782
 ; Sequence 1782, Application US/09764847
 ; Patent No. US20020132767A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC009
 ; CURRENT APPLICATION NUMBER: US/09/764,847
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PAMM or file wrapper
 ; NUMBER OF SEQ ID NOS: 2003
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1782
 ; LENGTH: 2651
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-847-1782

Query Match 57.6%; Score 577; DB 9; Length 2651;
 Best Local Similarity 100.0%; Pred. No. 1.0e-146;
 Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGAGCTGCACTGTTGTCACAGGACAGAGGCTGTGAGAGAGCTGCTGCTGCGAG 60
 Db 2075 ATGTGAGAGCTGCACTGTTGTCACAGGACAGAGGCTGTGAGAGAGCTGCTGCTGCGAG 2134
 QY 61 GACCTGCACTGCGGCTGTCTACTGTTGTGCTGCTGAGGCTGTGAGGAGGCTGCGAGT 120
 Db 2135 GACCTGCACTGCGGCTGTCTACTGTTGTGCTGCTGAGGCTGTGAGGAGGCTGCGAGT 2194
 QY 121 GGCCTGTGCTAACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 2195 GGCCTGTGCTAACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2254

Wed Dec 15 11:04:29 2004

us-09-995-225b-1.rmpb

Page 12

Search completed: December 15, 2004, 02:53:08
Job time : 514.891 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:17:59 ; Search time 3155.64 Seconds
(without alignments)
11570.599 Million cell updates/sec

Title: US-09-995-225B-1

Perfect score: 1002
Sequence: 1 atgtgagctgcagctgcgtc.....tgcagcagctgcgtgcgtag 1002

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : Listing first 45 summaries

EST: *
1: gb_ests1.*
2: gb_ests2.*
3: gb_nc.*
4: gb_ests3.*
5: gb_ests4.*
6: gb_ests5.*
7: gb_ests6.*
8: gb_g881.*
9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1002	100.0	1002	9 AY403440	AY403440 Homo sapi
2	1002	100.0	1787	3 CR605880	CR605880 full-length
3	1002	100.0	1809	3 CR619529	CR619529 full-length
4	1002	100.0	1861	3 CR619773	CR619773 full-length
5	883.8	88.2	1094	5 BX422717	BX422717 BX422717
6	843.6	84.2	1042	5 BX448207	BX448207 BX448207
7	839.2	83.8	1084	5 BX463176	BX463176 BX463176
8	757	75.5	1001	5 BX338900	BX338900 BX338900
9	742.8	74.1	1023	5 BX382511	BX382511 BX382511
10	665	66.4	907	5 BX332385	BX332385 BX332385
11	653.8	65.2	670	9 AY403441	AY403441 Pan trogl
12	647	64.6	851	4 B1825260	B1825260 Mus muscu
13	645.4	64.4	996	5 BQ712494	BQ712494 AGENCOURT
14	642	64.1	985	4 B1818819	B1818819 603037676
15	610	60.9	996	9 AY403442	AY403442 Mus muscu
16	610	60.9	1737	3 AKO89232	AKO89232 Mus muscu
17	610	60.9	1758	3 AKO46512	AKO46512 Mus muscu
18	608.4	60.7	1428	4 AKO87100	AKO87100 Mus muscu
19	598	59.7	773	4 B1910046	B1910046 603068162
20	596	59.5	1429	3 AKO50034	AKO50034 Mus muscu
21	592.2	59.1	925	4 BG326599	BG326599 602425455
22	586.2	58.5	940	4 B1754831	B1754831 603025292
23	581	58.0	709	4 B1820327	B1820327 603035201
24	549.8	54.9	826	4 B1818174	B1818174 603032623

25	542.4	54.1	544	4 BG570221	BG570221 602590622
26	539.8	53.9	650	2 BE731686	BE731686 601567082
27	512.6	51.2	914	5 BX377255	BX377255 BX377255
28	502	50.1	609	6 CD676083	CD676083 B81108.Y
29	501.6	50.1	824	4 B1767814	B1767814 603060921
30	467.6	46.7	699	4 BM007558	BM007558 603616944
31	443	44.3	1165	4 BM547338	BM547338 AGENCOURT
32	431.4	43.1	733	4 B1489911	B1489911 603031665
33	398.6	39.8	773	6 CB990075	CB990075 AGENCOURT
34	393.6	39.3	703	4 B181918	B181918 60303463
35	390	38.9	635	7 CF738689	CF738689 UI-M-HD0-
36	376.4	37.6	657	8 AZ801707	AZ801707 2M0060G16
37	372.4	37.2	697	7 CK836653	CK836653 4061701 B
38	364.6	36.4	557	6 BY747823	BY747823 BY747823
39	364	36.3	597	1 AA189791	AA189791 NC91C04.X
40	364	36.3	665	2 BB653568	BB653568 BB653568
41	360	35.9	499	5 BX488779	BX488779 DKFZ6860
42	353	35.2	692	4 BJ735246	BJ735246 BJ735246
43	352.8	35.2	886	7 CR587202	CR587202 CR587202
44	351.2	35.0	629	2 BB664126	BB664126 BB664126
45	349.4	34.9	808	5 BP711583	BP711583 BP711583

ALIGNMENTS

RESULT 1	AY403440	1002 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY403440				
DEFINITION	Homo sapiens HCM1558 gene, VIRUAL TRANSCRIPT, partial sequence.				
ACCESSION	AY403440				
VERSION	AY403440.1	GI:39759423			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1002)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Clivello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
REFERENCE	2 (bases 1 to 1002)				
AUTHORS	Todd,M.A., Tanenbaum,D.M., Clivello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
source	1..1002				
gene	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	<1..>1002				
	/locus_tag="HCM1558"				
ORIGIN					
Query Match	100.0%;	Score 1002;	DB 9;	Length 1002;	
Best Local Similarity	100.0%;	Pred. No. 1,2e-213;			
Matches 1002;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATGTGAGCTGCAGCTGTTCAACGACAGGAGCTGTGAGAGAGCTGCTGCTCCAG	60		
Db	1	ATGTGAGCTGCAGCTGTTCAACGACAGGAGCTGTGAGAGAGCTGCTGCTCCAG	60		

```

QY      61 GACCTGACGCTGGGGGCTGTCACCTGTTGCTGCTGGGCTGCTGGTGGGCTGTCAGT 120
Db      61 GACCTGACGCTGGGGGCTGTCACCTGTTGCTGCTGGGCTGCTGGTGGGCTGTCAGT 120
QY      121 GGCCTGTGCTAACACGCTGCTGTGTGCTGCTGACCACTACACAGCAAGGCCAGCATGACC 180
Db      121 GGCCTGTGCTAACACGCTGCTGTGTGCTGCTGACCACTACACAGCAAGGCCAGCATGACC 180
QY      181 ATGCCGAGCTGATCTTGTGCAACATGCGACATGCGACGCTGCTGCTCAGCGCCCTGACC 240
Db      181 ATGCCGAGCTGATCTTGTGCAACATGCGACATGCGACGCTGCTGCTCAGCGCCCTGACC 240
QY      241 CCTGTGACCTGCTGCGCCGCCGAGCTCCCGATGGGGGCTGTGAGTGGGGGGCGAA 300
Db      241 CCTGTGACCTGCTGCGCCGCCGAGCTCCCGATGGGGGCTGTGAGTGGGGGGCGAA 300
QY      301 GTCCACGTGGCACTGCAATCCCTTCAATGTCTCTGATGATGATGATGATGATGATGATGAT 360
Db      301 GTCCACGTGGCACTGCAATCCCTTCAATGTCTCTGATGATGATGATGATGATGATGATGAT 360
QY      361 GCCCTGTGAGCTGACCACTACATGAGGCTGACCTGACCTGACCTGACCTGACCTGACCT 420
Db      361 GCCCTGTGAGCTGACCACTACATGAGGCTGACCTGACCTGACCTGACCTGACCTGACCT 420
QY      421 GTGACAAACAGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db      421 GTGACAAACAGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY      481 TCCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db      481 TCCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY      541 ATGCAGAACGAGAGCTGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACG 600
Db      541 ATGCAGAACGAGAGCTGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACG 600
QY      601 CTGCGCACCTCTTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db      601 CTGCGCACCTCTTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY      661 GACCGGAGACAGCGGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCG 720
Db      661 GACCGGAGACAGCGGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCG 720
QY      721 AGCGATTTGGGCTCTGAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      721 AGCGATTTGGGCTCTGAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY      781 TCAGGAGGAGAGCGCGGCTGAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db      781 TCAGGAGGAGAGCGCGGCTGAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY      841 TCAGGAGGAGAGCGCGGCTGAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db      841 TCAGGAGGAGAGCGCGGCTGAGCGCACTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY      901 CAGAGCTTCCCGAGCAAGCTCCAGCGGCTGAGCAAGCTCCAGCGGCTGAGCAAGCTCCAG 960
Db      901 CAGAGCTTCCCGAGCAAGCTCCAGCGGCTGAGCAAGCTCCAGCGGCTGAGCAAGCTCCAG 960
QY      961 TCTCTCCCGAGCAAGCTCCAGCGGCTGAGCAAGCTCCAGCGGCTGAGCAAGCTCCAG 1002
Db      961 TCTCTCCCGAGCAAGCTCCAGCGGCTGAGCAAGCTCCAGCGGCTGAGCAAGCTCCAG 1002

```

```

RESULT 2
LOCUS   CR605880               1787 bp    mRNA       linear   HTC 21-Jul-2004
DEFINITION
full-length cDNA clone CS0DN01Y1C5 of Adult brain of Homo sapiens
ACCESSION
CR605880
VERSION CR605880.1  GI:50486687

```

```

KEYWORDS
SOURCE  HTC; CNSLT cDNA.
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE   1 (bases 1 to 1787)
JOURNAL Il,W.B., Gruber,C., Jesse,J., and Polayes,D.
REMARK  Full-length cDNA libraries and normalization
        Unpublished
        Contact : Feng Liang Email : fliang@life.techn.com URL :
        http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
        Faraday Avenue
        2 (bases 1 to 1787)
REFERENCE
AUTHORS Genoscope.
TITLE   Direct Submission
JOURNAL Submitted (20-Jul-2004) Genoscope - Centre National de Sequencage :
        BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)
        - Web : www.genoscope.cns.fr
COMMENT  1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime
        end enriched, double-strand cDNA was digested with Not I and cloned
        into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
        was normalized. Library was constructed by Life Technologies, a
        division of Invitrogen.
FEATURES
        Location/Qualifiers
            source
                1..1787
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DN01Y1C5"
                /tissue_type="Adult brain"
                /plasmid="pCMVSPORT_6"
ORIGIN
        Query Match      100.0%; Score 1002; DB 3; Length 1787;
        Best Local Similarity 100.0%; Pred. No. 1,2e-213;
        Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Oy		541	ATGCAAGACGGAGAAGCTGCAGCGCACACGCTGCTTCAATCCGCTACTAGTGTCACACA	600
Dd		611	ATGCAAAAGCAGAAAGCTGCCGACCACACGCTGGTTATCGCTACTAGTGTCACAGA	670
Oy		601	CTGGCCACCTCTCTACGCGCTGGTGTACTCTCCCGCTCCGACGAGGAAGAACGCCCCCTG	660
Dd		671	CTGGGCCA CCTCTACGCGCTGGTGTACTCTCCCGCTCCGACGAGGAAGAACGCCCCCTG	730
Oy		661	GACCGGGACACGGGCGCGGCTGGAGCGCTTCGGACACAGGCTGTCTGTGGCCACGCTGTGC	720
Dd		731	GACCGGGACACGGGCGCGGCTGGAGCGCTTCGGACACAGGCTGTCTGTGGCCACGCTGTGC	790
Oy		721	ACGCAATTGGGCTCTGGACGCGCACTATCTGATCTCTGTGGGACACAGGCTCATCTATC	780
Dd		791	ACGCAATTGGGCTCTGGACGCGCACTATCTGATCTCTGTGGGACACAGGCTCATCTATC	850
Oy		781	TCCGAGGGAAGACCCGCTGGACGCGCACTACTACTCTGGGCTACTGCACTTTGTGAAGATTTC	840
Dd		851	TCCGAGGGAAGACCCGCTGGACGCGCACTACTACTCTGGGCTACTGCACTTTGTGAAGATTTC	910
Oy		841	TCCAAATCTCCGGGCTTCTCCAGCAGCTTTTGA CACGACTTCTTACCGCTACATGAC	900
Dd		911	TCCAAATCTCCGGGCTTCTCCAGCAGCTTTTGA CACGACTTCTTACCGCTACATGAC	970
Oy		901	CAGAGCTTCCCACGACAGCTCCAACGCGTGTGAAAAAGCTGCCCTCGGGGACCCGAC	960
Dd		971	CAGAGCTTCCCACGACAGCTCCAACGCGTGTGAAAAAGCTGCCCTCGGGGACCCGAC	1030
Oy		961	TGCTTCCCCGACACATATGGGGGTGCAGACAGTGTCTGTGGCTAG	1002
Dd		1031	TGCTTCCCCGACACATATGGGGGTGCAGACAGTGTCTGTGGCTAG	1072
RESULT 3				
CR619529				
LOCUS				
DEFINITION	human-length cDNA clone CS0DM006YPO8 of fetal liver of Homo sapiens (human).	1809 bp	mRNA	linear HTC 21-JUL-2004
ACCESSION	CR619529			
VERSION	CR619529.1	GI:50500336		
KEYWORDS	HTC; CNSLT cDNA.			
SOURCE	Homo sapientia (human)			
ORGANISM	Homo sapientia			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Ll.M.B., Gruber,C., Jesssee,J. and Polayars,D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished			
REMARK	Contact : Feng liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/invitrogen Corporation 1600 Faraday Avenue Genoscope. 2 (baaeas 1 to 1809) Direct Submission Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five primes end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies, a division of invitrogen. Location/Qualifiers 1..1809 /organism="Homo sapientia" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DM006YPO8" /tissue_type="Fetal liver" /plasmid="pCMVSPORT_6"			
COMMENT				
FEATURES				
source				

[illegible]

[illegible]

Db	544	GTGTACACACGGGGACGTTGCGGCTTGCTGGGGTGGCGCGCTGCTGACCAAGTTT	603
Qy	481	TCTTCGTGCTCTTTTACATCTTGACCCATGTGTCCACCGGCGGTAGATGCGCAAG	540
Db	604	TCTTCCTCTCTCTTTACATCTTGACCCATGTGTCCACCGGCGCGCTAGATGCGCAAG	663
Qy	541	ATGCAGAACGCAAGACCTGCGACGCGCAGCGTGTGTATATGGCTACGTGGTCCAGCA	600
Db	664	ATGCAGAACGCAAGACCTGCGACGCGCAGCGTGTGTATATGGCTACGTGGTCCAGCA	723
Qy	601	CTGGCCACCTCTTACGCGCTGTGTCTATCTCCCGGTCCGACGAGGACACGCCCTG	660
Db	724	CTGGCCACCTCTTACGCGCTGTGTCTATCTCCCGGTCCGACGAGGACACGCCCTG	783
Qy	661	GACCGGGANACGGGGCGGCTGGAGCCCTCGGCACACAGGCTGCTGGTGGCCACCGTGTG	720
Db	784	GACCGGGANACGGGGCGGCTGGAGCCCTCGGCACACAGGCTGCTGGTGGCCACCGTGTG	843
Qy	721	ACGCAAGTTTGGGCTCTGGACGCGCACACTATCGATCTGTGGGGCACACGGTCAATCATC	780
Db	844	ACGCAAGTTTGGGCTCTGGACGCGCACACTATCGATCTGTGGGGCACACGGTCAATCATC	903
Qy	781	TGCGCAGGGAGCGCCGTGGACGCGCACACTACCTGGGGCTATGCACTTTGTGAAGGATTC	840
Db	904	TGCGCAGGGAGCGCCGTGGACGCGCACACTACCTGGGGCTATGCACTTTGTGAAGGATTC	963
Qy	841	TCCAAACTCTCGGCGCTTCTCCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC	900
Db	964	TCCAAACTCTCGGCGCTTCTCCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC	1023
Qy	901	CAGAGCTTCCCGACGACATCTCAAACGCGTGTGAAAAAGCTGCGCTGGCGGACCGGGAC	960
Db	1024	CAGAGCTTCCCGACGACATCTCAAACGCGTGTGAAAAAGCTGCGCTGGCGGACCGGGAC	1083
Qy	961	TGCTTCCCGGACCAATATGGGGGTGACGACAGGTGCTGGCGGTAG	1002
Db	1084	TGCTTCCCGGACCAATATGGGGGTGACGACAGGTGCTGGCGGTAG	1125

```

RESULT 5
BX422717
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX422717
BX422717 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
CSODN0011YC15 5-PRIME, mRNA sequence.
BX422717
BX422717.2 GI:46927958
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1094)
Li, W.B., Gruber, C., Jeese, J. and Pelayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:3064322.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4338.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODN001AB080P1&c=4338.f.
Location/Qualifiers
1..1094
/morganism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="GS0DN01YC15"
/issue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

ORIGIN

```

Query Match      88.2%; Score 883.8; DB 5; Length 1094;
Best Local Similarity 95.4%; Pred. No. 3,4e-187;
Matches 949; Conservative 13; Mismatches 25; Indels 8; Gaps 5;

QY 1 ATGTGAGCTGACAGCTGGTTCAACGCGACAGGGCTGTGAGAGAGCTGCTGCTGCGAG 60
DB ATGTGAGCTGACAGCTGGTTCAACGCGACAGGGCTGTGAGAGAGCTGCTGCTGCGAG 129
QY 61 GACCTGCACTGGGGCTGTCACTGTGTCCCTGCTGAGGCTGTGTGTGGGCTGTGCGAGT 120
DB 130 GACCTGCACTGGGGCTGTCACTGTGTCCCTGCTGAGGCTGTGTGTGGGCTGTGCGAGT 189
QY 121 GGCCTGTGTACAAAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 190 GGCCTGTGTACAAAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 249
QY 181 ATGCCGAGCGTGTACTTTGTCAACATGAGCAGTGTGAGGCTGTGTGTGTGTGTGTGTGT 240
DB 250 ATGCCGAGCGTGTACTTTGTCAACATGAGCAGTGTGAGGCTGTGTGTGTGTGTGTGTGT 309
QY 241 CCGTGTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
DB 310 CCGTGTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 369
QY 301 GTCCAGTGTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
DB 370 GTCCAGTGTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 429
QY 361 GGCCTGTGAGCTGTGACCACTAATGAGCAGTGTGAGGCTGTGTGTGTGTGTGTGTGTGT 420
DB 430 GGCCTGTGAGCTGTGACCACTAATGAGCAGTGTGAGGCTGTGTGTGTGTGTGTGTGTGT 489
QY 421 GTGTACAAACGCGGCGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB 490 GTGTACAAACGCGGCGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 549
QY 481 TCCCTGCTGTCTTCAATCTGACAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
DB 550 TCCCTGCTGTCTTCAATCTGACAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
QY 541 ATGCAGAACGACAGAGCTGTGCGACGCGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
DB 610 ATGCAGAACGACAGAGCTGTGCGACGCGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 669
QY 601 CTGGCCACCTCTTACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 659
DB 670 CTGGCCACCTCTTACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 729
QY 660 GGAACCGGAGACAGGGGCTGTGAGAGGCTGTGAGGCAACAGGCTGTGTGTGTGTGTGTGTGT 719
DB 730 GGAACCGGAGACAGGGGCTGTGAGAGGCTGTGAGGCAACAGGCTGTGTGTGTGTGTGTGTGT 789
QY 720 CACGCAAGTTT- GGGCTGTGAGCGCCACATATCTGATCTGTGTGTGTGTGTGTGTGTGTGT 778
DB 790 CACGCAAGTTT- GGGGCTGTGAGCGCCACATATCTGATCTGTGTGTGTGTGTGTGTGTGTGT 849
QY 779 TCTGTGAGAGGAGAGCCGCTGTGAGCGACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 838
DB 850 TCTGTGAGAGGAGAGCCGCTGTGAGCGACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 909
QY 839 TCTTCAAACTCTGTGGCTTCTCAGAGAGCTTTGTGACACCACTTCTTACCGGTACATGA 898

```

```

|||||
DB 910 TCTTCAAACTCTGTGGCTTCTCAGAGAGCTTTGTGAACCACTTCTTACCGGTACATGA 969
QY 899 ACCAGAGCTTCCCGACAGCTCCACGCTGTATGAAAAAGCTTCCCTGTGGGAGACCGGC 958
DB 970 AC--AAGAGCTTCCCGACAGAGC-YCAACGGCTGTATGAAAAAGHKCCCTGTGGGAG---CSGS 1023
QY 959 ACTGCTCCCGGACCAATGAGGAGTGTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 993
DB 1024 TMTTCTCCCGGACCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1058

```

RESULT 6

```

BX448207      1042 bp      mRNA      linear      EST 06-MAY-2004
LOCUS         BX448207 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION   CS0DM006YPO8 5-PRIME, mRNA sequence.
ACCESSION    BX448207
VERSION      BX448207.2 GI:47071231
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 1042)
AUTHORS      L.J.W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLES       Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT       On May 22, 2003 this sequence version replaced gi:31033927.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4338.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0AM006HD40P1ac=4338.f.
Location/Qualifiers
1..1042
source

```

FEATURES

```

source
1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM006YPO8"
/issue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

ORIGIN

```

Query Match      84.2%; Score 843.6; DB 5; Length 1042;
Best Local Similarity 97.0%; Pred. No. 3,3e-178;
Matches 883; Conservative 7; Mismatches 17; Indels 3; Gaps 3;

QY 1 ATGTGAGCTGACAGCTGGTTCAACGCGACAGGGCTGTGAGAGAGCTGCTGCTGCGAG 60
DB 76 ATGTGAGCTGACAGCTGGTTCAACGCGACAGGGCT- GTGAGAGAGCTGCTGCTGCGAG 134
QY 61 GACCTGCACTGGGGCTGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
DB 135 GACCTGCACTGGGGCTGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 194
QY 121 GGCCTGTGTACAAAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 195 GGCCTGTGTACAAAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 254

```

```

OY 181 ATGCCGACGCTGACTTTTGTCAACATGGCAGTGGCCTGTGCTCAGAGCCCTTGACC 240
DB 255 ATGCCGACGCTGACTTTTGTCAACATGGCAGTGGCCTGTGCTCAGAGCCCTTGACC 314
OY 241 CCTGTGACCTGTGCTGAGCCCTTCCAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGCGAA 300
DB 315 CTTGTGACCTGTGCTGAGCCCTTCCAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGCGAA 374
OY 301 GTCCACGTGGCACTGACATGCCATCCCTTCAATGTGTCTTCACTGTGGCCATGTACTCCACC 360
DB 375 GTCCACGTGGCACTGACATGCCATCCCTTCAATGTGTCTTCACTGTGGCCATGTACTCCACC 434
OY 361 GCCCTGTGAGCTTGCACCACTACATGAGCGTGCATGCGCGGAGCTACATAGGCCAAGC 420
DB 435 GCCCTGTGAGCTTGCACCACTACATGAGCGTGCATGCGCGGAGCTACATAGGCCAAGC 494
OY 421 GTGTACAACACGCGGACGCTGTGCGGCTTGTGTGGGGTGGCGGCTGTGACCACTTTC 480
DB 495 GTGTACAACACGCGGACGCTGTGCGGCTTGTGTGGGGTGGCGGCTGTGACCACTTTC 554
OY 481 TCCCTGTGCTCTTCTACATCTGTGAGGCACTGTGTCCACCCCGGCGGCTACAGTGGCGCAAG 540
DB 555 TCCCTGTGCTCTTCTACATCTGTGAGGCACTGTGTCCACCCCGGCGGCTACAGTGGCGCAAG 614
OY 541 ATGCAAGAACGAGAGGCTGCGGACGCGGACGCTGTGTTCATGCGCTACGTGTGCGGCA 600
DB 615 ATGCAAGAACGAGAGGCTGCGGACGCGGACGCTGTGTTCATGCGCTACGTGTGCGGCA 674
OY 601 CTGGCCACCTCTTACGCGCTGTGTCTACTCTCCCGGCTCCGACGAGGAGACAGGCCCTTC 660
DB 675 CTGGCCACCTCTTACGCGCTGTGTCTACTCTCCCGGCTCCGACGAGGAGACAGGCCCTTC 734
OY 661 GACCGGAGCAAGGGGCGGCTGGAGGCGCTCGGACACAGGCTGTGTGGCCACCGTGTGC 720
DB 735 GACCGGAGCAAGGGGCGGCTGGAGGCGCTCGGACACAGGCTGTGTGGCCACCGTGTGC 794
OY 721 AGCGAGTTGGGCTCTGTGAGCGCCACACTATGTATCTCTGTGGGCGACAGGTATATC 780
DB 795 AGCGAGTTGGGCTCTGTGAGCGCCACACTATGTATCTCTGTGGGCGACAGGTATATC 853
OY 781 TCGCAGAGGAGAGCCCGGTGAGCGCACACTACCTGGGGGCTACTGCACTTTGTGAAGATTC 840
DB 854 TCGCAGAGGAGAGCCCGGTGAGCGCACACTACCTGGGGGCTACTGCACTTTGTGAAGATTC 913
OY 841 TCCAAACTCTCTGCTCTTCCAGCAGCTTGTGACACCACTTCTTCACTCGCTACATGAAAC 900
DB 914 TCCAAACTCTCTGCTCTTCCAGCAGCTTGTGACACCACTTCTTCACTCGCTACATGAAAC 972
OY 901 CAGAGCTTCC 910
DB 973 AGAAGCTTCCC 982

```

RESULT 7
BX463176 1084 bp mRNA linear EST 06-MAY-2004
LOCUS BX463176 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM007YN21 5-PRIME, mRNA sequence.
ACCESSION BX463176
VERSION BX463176 GI:47069010
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1084)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31033492.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4338.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DM007CG110Pic=4338.f.
Location/Qualifiers
1..1084
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM007YN21"
/issue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN

Query Match 83.8%; Score 839.2; DB 5; Length 1084;
Best Local Similarity 97.3%; Pred. No. 3,2e-177;
Matches 858; Conservative 5; Mismatches 18; Indels 1; Gaps 1;

```

OY 1 ATGTGAGACTCAGCTGTTTCAACGGCAAGGCTGTGAGAGCTGCTGCTGCTGACG 60
DB 123 ATGTGAGACTCAGCTGTTTCAACGGCAAGGCTGTGAGAGCTGCTGCTGCTGACG 182
OY 61 GACCTGACGCTGGGCTGTCACTGTGTGCTGTGCGGCTGTGTGTGGCGTGGCAAGT 120
DB 183 GACCTGACGCTGGGCTGTCACTGTGTGCTGTGCGGCTGTGTGTGGCGTGGCAAGT 242
OY 121 GGCCTGTGCTCAACGCGCTGCTGTGTGCTGCGCACTTACACAGCAAGCCAGATGAC 180
DB 243 GGCCTGTGCTCAACGCGCTGCTGTGTGCTGCGCACTTACACAGCAAGCCAGATGAC 302
OY 181 ATGCCGACGCTGACTTTTGTCAACATGGCAGTGGCCTGTGCTCAGAGCCCTTGACC 240
DB 303 ATGCCGACGCTGACTTTTGTCAACATGGCAGTGGCCTGTGCTCAGAGCCCTTGACC 362
OY 241 CCTGTGACCTGTGCTGAGCCCTTCCAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGCGAA 300
DB 363 CCTGTGACCTGTGCTGAGCCCTTCCAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGCGAA 422
OY 301 GTCCACGTGGCACTGACATGCCATCCCTTCAATGTGTCTTCACTGTGGCCATGTACTCCACC 360
DB 423 GTCCACGTGGCACTGACATGCCATCCCTTCAATGTGTCTTCACTGTGGCCATGTACTCCACC 482
OY 361 GCCCTGTGAGCTTGCACCACTACATGAGCGTGCATGCGCGGAGCTACATAGGCCAAGC 420
DB 483 GCCCTGTGAGCTTGCACCACTACATGAGCGTGCATGCGCGGAGCTACATAGGCCAAGC 542
OY 421 GTGTACAACACGCGGACGCTGTGCGGCTTGTGTGGGGTGGCGGCTGTGACCACTTTC 480
DB 543 GTGTACAACACGCGGACGCTGTGCGGCTTGTGTGGGGTGGCGGCTGTGACCACTTTC 602
OY 480 CTCTGTGCTGCTTCTTCACTGTGAGCGCATGTGTCCACCGCGGCTGTGAGTGGCGCA 539
DB 603 CTCTGTGCTGCTTCTTCACTGTGAGCGCATGTGTCCACCGCGGCTGTGAGTGGCGCA 662
OY 540 GATGAGAGAGGAGAGGCTGCCACGCGCTGTGTGTTCATCTGCTACGTGTGTGCGAC 599
DB 663 GATGAGAGAGGAGAGGCTGCCACGCGCTGTGTGTTCATCTGCTACGTGTGTGCGAC 722
OY 600 ACTGGCCACCTCTACGCGCTGTGTGTCTTCTCCGCGTCCGAGGAGAGCAAGCGCCCT 659
DB 723 ACTGGCCACCTCTACGCGCTGTGTGTCTTCTCCGCGTCCGAGGAGAGCAAGCGCCCT 782

```

QY 660 GGACCGGAGACACGGGCGGCTGGAGCCCTCGAGCACAGGCTGCTGGTGCCACCGGTG 719
| | | | |
Db 783 GGACCGGAGACACGGGCGGCTGGAGCCCTCGAGCACAGGCTGCTGGTGCCACCGGTG 842
| | | | |
QY 720 CACGAGTTTGGGCTGTGAGCGCACATATCTGATCTGCTGGGGGACACGGTCAATCAT 779
| | | | |
Db 843 CACGAGTTTGGGCTGTGAGCGCACATATCTGATCTGCTGGGGGACACGGTCAATCAT 902
| | | | |
QY 780 CTGCGAGAGGAGCCCGTGGAGCACATACCTGCGGCTACTGCACTTTGTGAAGATTT 839
| | | | |
Db 903 CTGCGAGAGGAGCCCGTGGAGCACATACCTGCGGCTACTGCACTTTGTGAAGATTT 962
| | | | |
QY 840 CTCGAACCTCTGCGCCCTTCTCGAGCAGCTTTGTGACACCACT 881
| | | | |
Db 963 CTCGAACCTCTGCGCCCTTCTCGASMRCTTTGTGAMACACTTCT 1004
| | | | |
RESULT 8
BX38900 1001 bp mRNA linear EST 07-APR-2004
LOCUS BX38900 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION Clons CS0D1065YE04 5-PRIME, mRNA sequence.
ACCESSION BX38900
VERSION BX38900.2 GI:46268457
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1001)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayer, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 2, 2003 this sequence version replaced gi:30333744.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4338.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D1065BC020P1ac=4338.f.
FEATURES
source
1..1001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1065YE04"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized."
ORIGIN
Query Match 75.5%; Score 757; DB 5; Length 1001;
Best Local Similarity 95.2%; Pred. No. 8e-159;
Matches 789; Conservative 20; Mismatches 15; Indels 5; Gaps 3;
QY 1 ATGTGAGCTGACGTGTTCAACGACAGGCGTGTGAGAGCTGCTGCTGCAG 60
| | | | |
Db 129 ATGTGAGCTGACGTGTTCAACGACAGGCGTGTGAGAGCTGCTGCTGCAG 188
| | | | |
QY 61 GACCTGACGCTGGGCTGTCTACTGTGTGCTGCTGGGCTGTGTGTGGGCTGCCAGT 120
| | | | |
Db 189 GACCTGACGCTGGGCTGTCTACTGTGTGCTGCTGGGCTGTGTGTGGGCTGCCAGT 248
| | | | |
QY 121 GGCTGTGCTACAAAGCCCTGCTGTGTGCTGACCACTACACAGCAAGGCGACATGACC 180
| | | | |

Db 249 GGCTGTGCTACAAAGCCCTGCTGTGTGCTGACCACTACACAGCAAGGCGACATGACC 308
| | | | |
QY 181 ATGCCGACGTGTACTTTGTCAACATGACAGTGGCAGGCTGTGCTCACGCCCTTGCC 240
| | | | |
Db 309 ATGCCGACGTGTACTTTGTCAACATGACAGTGGCAGGCTGTGCTCACGCCCTTGCC 358
| | | | |
QY 241 CCTGTGACCTGTGCGGCCCGCCGAGCTCCCGTGGGCGCTGTGAGTGTGGGCGGAA 300
| | | | |
Db 369 CCGTGACCTGTGCGGCCCGCCGAGCTCCCGTGGGCGCTGTGAGTGTGGGCGGAA 428
| | | | |
QY 301 GTCCAGTGGACCTGACAGATCCCTTCAATGTGTCTCACTGTGCGCACTATACAC 360
| | | | |
Db 429 GTCCAGTGGACCTGACAGATCCCTTCAATGTGTCTCACTGTGCGCACTATACAC 488
| | | | |
QY 361 GCCCTGCTAGGCTCGACCACTACATGAGCTGCACTCCCGGACCTACATGCGCAGC 420
| | | | |
Db 489 GCCCTGCTAGGCTCGACCACTACATGAGCTGCACTCCCGGACCTACATGCGCAGC 548
| | | | |
QY 421 GTGTACAAACGGCGG-CAGGTGTGGGCTTGGTGGGGTGGGCGGCTGTGACAGCTT 479
| | | | |
Db 549 GTGTACAAACGGCGGACAGTGTGCGGCTTGTGTGGGGTGGGCGGCTGTGACAGCTT 608
| | | | |
QY 480 CTCCTGCTGCTCTTCTACATCTGACCATGTGTCCACCCGCGCTAGAGTGCAGCA 539
| | | | |
Db 609 CTCCTGCTGCTCTTCTACATCTGACCATGTGTCCACCCGCGCTAGAGTGCAGCA 668
| | | | |
QY 540 GATGCAAGACGAGAAAGCTGCCGACCGCAAGCTGTTCATCGCTAGTGTGCAGC 599
| | | | |
Db 669 GATGCAAGACGAGAAAGCTGCCGACCGCAAGCTGTTCATCGCTAGTGTGCAGC 728
| | | | |
QY 600 ACTGGCACCTCTTACGCGCTGTGTACTCTCCCGCGTCCGAGGAGGACACGCGCCCT 659
| | | | |
Db 729 ACTGGCACCTCTTACGCGCTGTGTACTCTCCCGCGTCCGAGGAGGACACGCGCCCT 788
| | | | |
QY 660 GGACCGGAGACACGGGCGGCTGGAGCCCTCGAGCACAGGCTGCTGATGCGCACCGTGTG 719
| | | | |
Db 783 GGACCGGAGACACGGGCGGCTGGAGCCCTCGAGCACAGGCTGCTGATGCGCACCGTGTG 848
| | | | |
QY 720 CACGAGTTTGGGCTGTGAGCGCACATATCTGATCTGCTGGGGGACACGGTCAATCAT 779
| | | | |
Db 849 CACGAGTTTGGGCTGTGAGCGCACATATCTGATCTGCTGGGGGACACGGTCAATCAT 904
| | | | |
QY 780 CTGCGAGAGGAGCCCGTGGAGCACATACCTGCGGCTACTGCACTTTGTGAAGATTT 828
| | | | |
Db 905 CTGCGAGAGGAGCCCGTGGAGCACATACCTGCGGCTACTGCACTTTGTGAAGATTT 953
| | | | |
RESULT 9
BX382511 1023 bp mRNA linear EST 23-APR-2004
LOCUS BX382511 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION Clons CS0D1081YB16 5-PRIME, mRNA sequence.
ACCESSION BX382511
VERSION BX382511.2 GI:46557694
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1023)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayer, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30439943.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster 4338.f

For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?c=CSOD1081DA08QPl&c=4338.f>.

FEATURES

source

1..1023
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1081YB1.6"
/issue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match

Best Local Similarity 74.1%; Score 742.8; DB 5; Length 1023;
Matches 818; Conservative 11; Mismatches 44; Indels 5; Gaps 5;

```
QY 1 ATGTGAGCTGAGCTGTTCAAGGCAAGGCGTGTGGAGAGAGCTGCTGCTCCAG 60
Db 98 ATGTGAGCGGAGCGGCTTCAACGCAAGGCTGTGGAGAGAGCGCTGCTGCTCCAG 157
QY 61 GACCTGCAAGCTGGGCTGTCACTGTTGTGCTGTGGCTGTGTGTGGGCTGCTGAG 120
Db 158 GACCTGCAAGCGGCGGCTGTCACTGTTGTGCTGTGGCTGTGTGTGGGCTGCTGAG 217
QY 121 GACCTGCTGCTCAACAGCGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGGCTGAG 180
Db 218 GACCTGCTGCTCAACAGCGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGGCTGAG 277
QY 181 ATGCGGAGCTGTACTTTGTCAATGAGTGTGCAAGGCTGTGTGTGGCTGTGTGTGG 240
Db 278 ATGCGGAGCTGTACTTTGTCAATGAGTGTGCAAGGCTGTGTGTGGCTGTGTGTGG 337
QY 241 CCTGTGCACTGCTGCGGCGGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGGCTGAG 300
Db 338 CCTGTGCACTGCTGCGGCGGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGGCTGAG 397
QY 301 GTCCAGCTGTGCAATGAGTGTGCAATGAGTGTGCAATGAGTGTGCAATGAGTGTGCA 360
Db 398 GTCCAGCTGTGCAATGAGTGTGCAATGAGTGTGCAATGAGTGTGCAATGAGTGTGCA 457
QY 361 GCGCTGTGAGCTGTGCAATGAGTGTGCAATGAGTGTGCAATGAGTGTGCAATGAGTGTG 420
Db 458 GCGCTGTGAGCTGTGCAATGAGTGTGCAATGAGTGTGCAATGAGTGTGCAATGAGTGTG 517
QY 421 GTGTACAAACGCGG-CACGTGTGCGGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGG 479
Db 518 GTGTACAAACGCGG-CACGTGTGCGGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGG 577
QY 480 CTGCTGCTGCTTTTATCATCTGCAAGCATGTGTGCAAGCATGTGTGCAAGCATGTGTG 539
Db 578 CTGCTGCTGCTTTTATCATCTGCAAGCATGTGTGCAAGCATGTGTGCAAGCATGTGTG 637
QY 540 GATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599
Db 638 GATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 697
QY 600 ATTGCGCACTCTTACGCGGCTGTGTGTGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGG 659
Db 698 ATTGCGCACTCTTACGCGGCTGTGTGTGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGG 757
QY 660 GAGACGGGAGCAAGGCGGCTGTGTGTGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGG 719
Db 758 GAGACGGGAGCAAGGCGGCTGTGTGTGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGG 817
QY 720 CACGCAATTGGGCTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 779
Db 818 CACGCAATTGGGCTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875
```

```
QY 780 CTGCGGAGGAGAGCGGCTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 839
Db 876 CTGCGGAGGAG-ARCCCGGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 933
QY 840 CTCGAACCTCTGAGCTTCTCCAGAGCTTTGTGACAC 877
Db 934 YTCGAACCTCTGAGCTTCTCCAGAGCTTTGTAACACTC 971
```

RESULT 10

EX332385 907 bp mRNA linear EST 07-APR-2004
LOCUS BX332385 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CSODC011YB12 5-PRIME, mRNA sequence.
ACCESSION BX332385
VERSION BX332385.2 GI:46264178
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 907)
Li W.B., Gruber,C., Jeesee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 1, 2003 this sequence version replaced gi:30310147.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4338.f

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

source

For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?c=CSODC011DA06QPl&c=4338.f>.
Location/Qualifiers
1..907
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODC011YB1.2"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match

Best Local Similarity 66.4%; Score 665; DB 5; Length 907;
Matches 745; Conservative 7; Mismatches 12; Indels 7; Gaps 7;

```
QY 1 ATGTGAGCTGAGCTGTTCAAGGCAAGGCGTGTGGAGAGAGCTGCTGCTCCAG 60
Db 113 ATGTGAGCTGAGCTGTTCAAGGCAAGGCGTGTGGAGAGAGCTGCTGCTCCAG 172
QY 61 GACCTGCAAGCTGGGCTGTCACTGTTGTGCTGTGGCTGTGTGTGGCTGTGTGTGG 120
Db 173 GACCTGCAAGCTGGGCTGTCACTGTTGTGCTGTGGCTGTGTGTGGCTGTGTGTGG 232
QY 121 GGCCTGTGCTCAACAGCGCTGTGTGTGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGG 180
Db 233 GGCCTGTGCTCAACAGCGCTGTGTGTGCTGTGTGTGGCTGTGTGTGGCTGTGTGTGG 292
QY 181 ATGCGGAGCTGTACTTTGTCAATGAGTGTGCAAGGCTGTGTGTGGCTGTGTGTGG 240
Db 293 ATGCGGAGCTGTACTTTGTCAATGAGTGTGCAAGGCTGTGTGTGGCTGTGTGTGG 352
```


QY 241 CTTGTGACC-TGCTGGGCCCCCGAGCTCCGGTGGGCGTGTGAGTGTGGGCGGCGA 299
 Db 353 CCGGTGACCATGCTGCGGCCCCCGAGCTCCGGTGGGCGTGTGAGTGTGGGCGGCGA 412
 QY 300 AGTTCACGTGGCACTGCAAGATCCCTTCAATGTGTCTCACTGTGGCCATGTACTGCAC 359
 Db 413 AGTTCACGTGGCACTGCAAGATCCCTTCAATGTGTCTCACTGTGGCCATGTACTGCAC 472
 QY 360 CGGCTGTGAGGCTGACCACTACATCAGCGCTGACTGCTCCGCGAAGCTT-ACATGGCCA 418
 Db 473 CGGCTGTGAGGCTGACCACTACATCAGCGCTGACTGCTCCGCGAAGCTT-ACATGGCCA 532
 QY 419 GCGTGTACACACAGCGGCACTGTGCGGCTTGTGTGGGTGGGCGGCTGTGACCAAGCT 478
 Db 533 GCGTGTACACACAGCGGCACTGTGCGGCTTGTGTGGGTGGGCGGCTGTGACCAAGCT 552
 QY 479 TCTCTGTGCTCTTCTACATGTGACAGCATGTGTGACCGCGGCTGTGAGTGTGCGCA 538
 Db 593 TCTCTGTGCTCTTCTACATGTGACAGCATGTGTGACCGCGGCTGTGAGTGTGCGCA 652
 QY 539 AGATGACAGAACGAGAAAGCTGCGGACGCGCTGTGTCTATCGGCTTACGTGTGCGAG 598
 Db 653 AGATGACAGAACGCG-RAARCTGCGGCGGCGGCTGTGTCTATCGGCTTACGTGTGCGAG 711
 QY 599 CACTGGCAACCTCTGTGCGGCTGTGTCTCTCCGCGTGTGCGGAGGAGAGACAGCGGCC 658
 Db 712 CACTGGCAACCTCTGTGCGGCTGTGTCTCTCCGCGTGTGCGGAGGAGAGACAGCGGCC 771
 QY 659 TGGACG-GGGACAGCGGCGGCTGTGAGCGCTGTGCGGACAGCAAGGCTGTGTGCGCA 717
 Db 772 TGGACGCGGAGACAGCGGCGGCTGTGAGCGCTGTGCGGACAGCAAGGCTGTGTGCG-CA 830
 QY 718 TGCACGCACTTTGGGCTCTGTGAGCGCACACTATCTGTCTGCTGGGCGCAC 768
 Db 831 TGCACGCA-CTTGGGCTGTGTGAGCGCC-GACTATCTGTATCTGCTGGGAGMAC 879

RESULT 11
 LOCUS AY403441 670 bp DNA linear GSS 12-DEC-2003
 DEFINITION Pan troglodytes HCM1558 gene, VIRUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY403441
 VERSION AY403441.1 GI:39759424
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 1 (bases 1 to 670)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 670)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submision
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1..670
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"

gene <1..>670
 ORIGIN /locus_tag="HCM1558"
 Query Match 65.2%; Score 653.8; DB 9; Length 670;
 Best Local Similarity 98.2%; Pred. No. 9, 6e-136;
 Matches 658; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 332 TGTCTCATCTGTGTGGGCAATGTAATCCACCGGCTGTGTGAGGCTTGGCACTATATGAGC 391
 Db 1 TGTCTCATCTGTGTGGGCAATGTAATCCACCGGCTGTGTGAGGCTTGGCACTATATGAGC 60
 QY 392 GTSCATGCGCGGAGACCTTACATGAGGAGGTGTAACAAGCGGCGAGTGTGGGCTTGG 451
 Db 61 GTSCATGCGCGGAGACCTTACATGAGGAGGTGTAACAAGCGGCGAGTGTGGGCTTGG 120
 QY 452 TGTGGGTGTGGCGGCTGTGACCAAGCTTCTCTGTGTCTTCTACATCTGAGGCTATG 511
 Db 121 TGTGGGTGTGGCGGCTGTGACCAAGCTTCTCTGTGTCTTCTACATCTGAGGCTATG 180
 QY 512 TGTTCACCGCGCGGCTGTGAGTGTGGGCAATGTAATGTAACAAGCTTGTGGGCGGAG 571
 Db 181 TGTTCACCGCGCGGCTGTGAGTGTGGGCAATGTAATGTAACAAGCTTGTGGGCGGAG 240
 QY 572 TGTGTTCATCGGCTGTGAGTGTGTCAGCACTGAGCAAGCTTGTGAGGCTGTGCTACT 631
 Db 241 TGTGTTCATCGGCTGTGAGTGTGTCAGCACTGAGCAAGCTTGTGAGGCTGTGCTACT 300
 QY 632 CCGGCTGTGCGGAGAGAGACAGCGCCCTGTGACCGGAGACAGGCGGCTGTGAGCCTTGG 691
 Db 301 CCGGCTGTGCGGAGAGAGACAGCGCCCTGTGACCGGAGACAGGCGGCTGTGAGCCTTGG 360
 QY 692 CACACAGGCTGTGTGTGGGCAAGCTTGTGAGCAAGATTTGGGCTGTGAGCGGCACTAT 751
 Db 361 CACACAGGCTGTGTGTGGGCAAGCTTGTGAGCAAGATTTGGGCTGTGAGCGGCACTAT 420
 QY 752 TGATCTGTGGGCGGAGACAGGCTATCATCTGCGAGGAGAGCCGTGTGAGCGACACTAC 811
 Db 421 TGATCTGTGGGCGGAGACAGGCTATCATCTGCGAGGAGAGCCGTGTGAGCGACACTAC 480
 QY 812 TGGGCTGTGCACTTGTGAGAGATTTGTCCAACTCTGTGAGGCTTGTGAGGAGCTTGG 871
 Db 481 TGGGCTGTGCACTTGTGAGAGATTTGTCCAACTCTGTGAGGCTTGTGAGGAGCTTGG 540
 QY 872 TGACACCACTTGTCTTACCGCTACATGACCAAGAGCTTCCGAGCAAGCTTCAAGCTGA 931
 Db 541 TGACACCACTTGTCTTACCGCTACATGACCAAGAGCTTCCGAGCAAGCTTCAAGCTGA 600
 QY 932 TGAAGAAAGCTGCGGCTGTGCGGAGACCGGCACTGCTCCCGAGACCAATGAGGAGTGCAGAG 991
 Db 601 TGAAGAAAGCTGCGGCTGTGCGGAGACCGGCACTGCTCCCGAGAGGAGTGCAGAGAG 660
 QY 992 TGCTGTGCGTA 1001
 Db 661 TGCTGTGCGTA 670

RESULT 12
 B1825260
 LOCUS B1825260 851 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603071957F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5163926 5',
 mRNA sequence.
 ACCESSION B1825260
 VERSION B1825260.1 GI:15936810
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 851)
 REFERENCES
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: csabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1M11407 row: a column: 15
 High quality sequence stop: 791.
 Location/Qualifiers

FEATURES
source

```
1..851
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5163926"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: Brain; Vector: pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."
```

ORIGIN

Query Match 64.4%; Score 647; DB 4; Length 851;

Best Local Similarity 96.4%; Pred. No. 3,3e-134; Indels 7; Gaps 7; Matches 727; Conservative 0; Mismatches 5;

```

QY 6 GAGCTGACGCTGTTCAACGCGACAGGCGCTGTGAGAGACTGCTGCTGCCAGACCT 65
DB 114 GAGCTGACGCTG-TCAACGCGACAGGCGCTGTGAGAGACTGCTGCTGCCAGACCT 172
QY 66 GCAGCTGGGCTGTCACTGTTGTGCTGTGGGCTGTGTGGGCTGTGCCAGTGGGCT 125
DB 173 GCAGCTGGGCTGTCACTGTTGTGCTGTGGGCTGTGTGGGCTGTGCCAGTGGGCT 232
QY 126 GTGCTACAAGCCGCTGTGTGCTGCGCAACTTACAGCAAGGCGCAGCATGACATGCC 185
DB 233 GTGCTACAAGCCGCTGTGTGCTGCGCAACTTACAGCAAGGCGCAGCATGACATGCC 232
QY 186 GGAAGTACTTTGTCAACATGAGGAGTGGAGGCTGTGCTAGCGCCCTGACCCCTGT 245
DB 293 GGAAGTACTTTGTCAACATGAGGAGTGGAGGCTGTGCTAGCGCCCTGACCCCTGT 352
QY 246 GCACCTGTGGGCGGCGGAGCTCCCGTGGGCGCTGTGAGTGTGGGCGGCGCAAGTCCA 305
DB 353 GCACCTGTGGGCGGCGGAGCTCCCGTGGGCGCTGTGAGTGTGGGCGGCGCAAGTCCA 412
QY 306 CGTGGCACTGAGATCCCTTCAATGTGTCTCACTGGTGGCATGTACTCAACGGCCT 365
DB 413 CGTGGCACTGAGATCCCTTCAATGTGTCTCACTGGTGGCATGTACTCAACGGCCT 472
QY 366 GCTGAGCTCGACACTACATGAGCGGTGACTGCCGCGAAGCTTACATGCGCAGCGTGA 425
DB 473 GCTGAGCTCGACACTACATGAGCGGTGACTGCCGCGAAGCTTACATGCGCAGCGTGA 532
QY 426 CAACACGGGCACTGTGCGGCTGTGTGGGCTGTGGGCGGCGTGTG-ACGAGCTTCTCT 484
DB 533 CAACACGGGCACTGTGCGGCTGTGTGGGCTGTGGGCGGCGTGTG-ACGAGCTTCTCT 592
QY 485 CGCTGCTCTTCTACATCTGACGCGATGTGTCCACCGCG-GGCTAGAGTGGCGCAAGTGG 543
DB 593 CGCTGCTCTTCTACATCTGACGCGATGTGTCCACCGCGAGGCTAGAGTGGCGCAAGTGG 652
QY 544 CAGAACGCAAGCTGCCGACCGACGCTGTGTTCATGCGCTACGTGTGCCAGCACTG 603

```

DB 653 CAGAACGCAAGCTGCCGACCGCAGCTGTGTTCATGCGCTACGTGTGCCAGCACTG 712

QY 604 GCCACCTCTACGAGCTGTGTGCTACTCTCCGCGGTCCGAGAGAGACAGCCCTGAGAC 663

DB 713 GCCACCTCTACGAGCTGTGTGCTACTCTCCGCGGTCCGAGAGAGAGACAGCCCTGAGAA 772

QY 664 CGGACACGCGGCGGCTGTGGA-GCCCTGGCAC-ACAGGCTGTGG-TGGCCACGGTGTG 719

DB 773 CGGACACGCGGCGGCTGTGAGAGCCCTGGCACACAGGCTGTGTGATGGCCACGCTGTGG 832

QY 720 CACGCACTTTGGGCTCTGG 738

DB 833 CACGCACTTTGGGCTCTGAG 851

RESULT 13

BQ712494 996 bp mRNA linear EST 16-JUN-2002

LOCUS AGENCOURT 8352557 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6277823

DEFINITION 5', mRNA sequence.

ACCESSION BQ712494.1 GI:21851393

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.

AUTHORS 1 (bases 1 to 996)

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

EMAIL: csabbs-remail.nih.gov

Tissue Procurement: Dr. Mark Watson

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: L1M2464 row: m column: 24

High quality sequence stop: 611.

Location/Qualifiers

1..996

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6277823"

/lab_host="DH10B (phage-resistant)"

/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 64.4%; Score 645.4; DB 5; Length 996;

Best Local Similarity 96.1%; Pred. No. 7,6e-134; Indels 2; Gaps 2; Matches 683; Conservative 0; Mismatches 26;

```

QY 1 ATGTGAGCTGACGCTGTTCACAGGCGACAGGCGTGTGAGAGCTGCTGCTGCCAG 60
DB 151 ATGTGAGCTGACGCTGTTCACAGGCGACAGGCGTGTGAGAGCTGCTGCTGCCAG 210
QY 61 GACCTGACGCTGGGCGCTGTCACTGTTGTGCTGTGGGCTGTGTGGTGGCGTGGCAAGT 120
DB 211 GACCTGACGCTGGGCGCTGTCACTGTTGTGCTGTGGGCTGTGTGGTGGCGTGGCAAGT 270
QY 121 GGCTGTGCTAACAGCCCTGCTGTGTGTGGCCAACTACAGCAAGGCGCATGAC 180

```



```

VERSION      AY403442.1  GI:39759425
KEYWORDS
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
REFERENCE    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS      1 (bases 1 to 996)
              Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejaritwal, A.,
              Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
              Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
              Adams, M.D. and Cargill, M.
              Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED       14671302
REFERENCE    2 (bases 1 to 996)
              Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejaritwal, A.,
              Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
              Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
              Adams, M.D. and Cargill, M.
              Direct Submision
              Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT      This sequence was made by sequencing genomic exons and ordering
              them based on alignment
FEATURES
  source      location/Qualifiers
              1..996
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10090"
              <1..>996
              /locus_tag="HCM155b"
ORIGIN
Query Match      60.9%; Score 610; DB 9; Length 996;
Beet Local Similarity 76.9%; Pred. No. 6.3e-126;
Matches 771; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

QY 1 ATGTGAGAGCTGACAGTGGTTCAACGACAGAGGCTGGTGAAGAGCTGCTGCTGCAG 60
DB 1 ATGTGAGAGCTGAGCCCACTCAACAGACAGCGTGGAGTGAAGCGC--TGTCGG 57
QY 61 GACCTGAGAGCTGGGGCTGTCACTGTTGTCCGCTGCGGCTGGTGGTGGCGCTGCAATG 120
DB 58 AACCTGCGCTGGGGCTGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 117
QY 121 GGCCTGTGCTTAAACGCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
DB 118 AGCTTAGAGCTATATATGCTCTTCTGAGCTGGCCACCTGAGCAGCAAGAACATGACC 177
QY 181 ATGCCGAGCGTGTACTTGTTCATCATGAGAGTGGAGAGGCTGTGCTCAAGCGCTGGCC 240
DB 178 ATGCCGAGCGTGTACTTGTTCATCATGAGAGTGGAGAGGCTGTGCTCAAGCGCTGGCC 237
QY 241 CCTGTGACCTGTGCGGCCCCCGGAGCTCCGCTGGGCGCTGTGAGTGGTGGCGGAG 300
DB 238 CTTGCTACTCTGTTGGGCGCTGCGCACTCAAGTGGCGCTGTGAGGCTCAAGCATGAG 297
QY 301 GTTCAAGTGGCATGATCCCTTCAATGATGCTCACTGATGCTGATGCTGATGCTGATGCT 360
DB 298 GCCCATGTGACATGCTGATCTCTTCAAGTGGCTTCCCTGTGATGATGATGATGATGAT 357
QY 361 GCCCTGTGAGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 358 GCACTGTGAGCTTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 417
QY 421 GTGTACAACAAGCGGAGTGGGCTTGTGAGTGGGCTTGTGAGTGGGCTTGTGAGTGGGCTT 480
DB 418 GTGTACAACAAGCGGAGTGGGCTTGTGAGTGGGCTTGTGAGTGGGCTTGTGAGTGGGCTT 477
QY 481 TCTTGTGCTGCTTCTTCACTGTGACGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 478 TCTTGTGCTGCTTCTTCACTGTGACGATGATGATGATGATGATGATGATGATGATGATGAT 537

```

```

QY 541 ATGTGAGAGCGAAGAGTGTCCGACGCGACGCTGTGTTTCAATGAGCTACGTGAGCGACA 600
DB 538 ATGTGAGAGACAGGAGGAGCGAGCTGATCTTGTGCTCATGAGGCTACAGTGTCCAGGT 597
QY 601 CTGGCCACCTCTTACCGCGCTGATGATCTCTCCGCTGCGAGGAGGAGACAGGCTCCGTG 660
DB 598 CTGGCTGTGTGTATATCCCTCTGACATCATCTGAGATTCGGAGAGAGACACACCTCTG 657
QY 661 GATCCGGAGACAGGCGCGCTGTGAGAGCTTGTGACACAGGCTGCTGTGTGGCCACGCTGTC 720
DB 658 GACCAAGACACAGACAGGCTGTGAGACCTCTGTGATGACAGGCTGTGTGTGTGTGTGTGT 717
QY 721 ACGCATTTGGGCTGTGAGACGCGACATCTGATCTGTGCTGTGGGACACAGGTATCATC 780
DB 718 ACTCATTTGGGCTGTGAGACGCGACATCTGATCTGTGAGC--CTGGGGACACAGTGTGACG 774
QY 781 TCGCAGAGGAGACCGCTGTGAGACGACATCACTGAGGCTGTGACCTTTGTGAGAGATTC 840
DB 775 TCAAGGGGAGAGACCGTGTGAGGAGATATCTGGCATCTTACAGTTGTGAGAGACCTG 834
QY 841 TCAAACTCTGGGCTTTCTTCAAGACGCTTTGTGACCACTTCTTACCGCTACATGAG 900
DB 835 GCGAGTTCTGGGCTTTCTTCAAGACGATCTGTGTGACCACTGTCTTACCGTTATCATAC 894
QY 901 CAGAGCTTCCCGAGAGCTCCAAAGGCTGATGAAAAAGCTGCCCTGGCGGGAGACCGGAC 960
DB 895 AAAGCTTCCCGGAGAGCTCCGCGGCTGTATGAGAGATGACCTGCGGGCGCGGAC 954
QY 961 TGCTCCCGGAGACCATGAGGAGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 1002
DB 955 TGCTCCCGGAGACCTCTGGGAGATGACAGAGTGTGAGAG 996

```

Search completed: December 15, 2004, 02:34:59
Job time : 3160.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 18:35:19 ; Search time 53.3965 Seconds
(without alignments)
2237.166 Million cell updates/sec

Title: US-09-995-225b-2

Perfect score: 1733
Sequence: 1 MWSGSMFNGTGLVEELPACQ.....LPCCDRHCSPPDHMGVQVLA 333

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_23sep04:*

1: geneseqp1980s:.*
2: geneseqp1980s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1733	100.0	333	4	AAU25607 Human G P
2	1733	100.0	333	5	ABJ04068 Human G P
3	1733	100.0	333	5	AAU97734 Human G P
4	1733	100.0	333	6	ABR56303 Human G-P
5	1733	100.0	333	6	ABU09900 Human G-P
6	1733	100.0	333	7	ADG65653 Human GPC
7	1733	100.0	333	8	ADO29606 Human GPC
8	1733	100.0	370	4	AAW25659 Human PRC
9	1733	100.0	695	6	ABR56305 pc901HISG
10	1733	100.0	714	6	ABR56304 pc901HISG
11	1733	100.0	735	6	ABR56306 pc901HISG
12	1726	99.6	333	3	AAV57285 Human GPC
13	1726	99.6	333	4	AAU25584 Human G P
14	1726	99.6	333	5	ABU54867 Human G-P
15	1717	99.1	333	3	AAV78856 Human che
16	1717	99.1	333	6	ABP81753 Human G P
17	1276	73.6	333	8	ADO29607 Mouse GPC
18	795	45.9	153	5	ABBO4632 Human ops
19	711	41.0	142	5	ABP10980 Human ORF
20	271	15.6	375	2	AAW03739 rchd523 g
21	271	15.6	375	2	AAW89302 Human rch
22	271	15.6	375	2	AAW95158 Amln ac1
23	271	15.6	375	2	AAW95158 Amln ac1
24	271	15.6	375	3	AAW78506 Protein s
25	271	15.6	375	3	AAW68445 Rchd523 a

26	271	15.6	375	7	AD110059	Ad110059 Novel hum
27	271	15.6	375	8	ADP91255	AdP91255 Human end
28	269	15.5	375	3	AAV90658	AAV90658 Human mut
29	269	15.5	375	3	AAV90662	AAV90662 Human mut
30	269	15.5	375	7	ADC22713	ADC22713 Human G P
31	269	15.5	375	7	ADC22721	ADC22721 Human G P
32	269	15.5	375	7	ADH14194	Adh14194 Mutated h
33	269	15.5	375	7	ADH14194	Adh14194 Mutated h
34	267	15.4	375	2	AAW22338	AAW22338 PEG-1 pro
35	267	15.4	375	3	AAV59177	AAV59177 Human CDP
36	267	15.4	375	3	AAV90628	AAV90628 Human G P
37	267	15.4	375	3	AAV90624	AAV90624 Human G P
38	267	15.4	375	6	ABG73996	ABG73996 Human CMK
39	267	15.4	375	6	ABU08546	ABU08546 Human hep
40	267	15.4	375	6	ABP81804	ABP81804 Human G P
41	267	15.4	375	7	ADC22589	ADC22589 Human G P
42	267	15.4	375	7	ADC22575	ADC22575 Human G P
43	267	15.4	375	7	ADH14048	Adh14048 Human GPR
44	267	15.4	375	7	ADH14062	Adh14062 Human CDP
45	267	15.4	375	8	ADO29390	ADO29390 Human GPC

ALIGNMENTS

RESULT 1	AAU25607	standard; protein; 333 AA.
ID	AAU25607	
XX	AAU25607;	
AC	18-DEC-2001 (first entry)	
XX		
DT		
DE	Human G Protein-Coupled Receptor (GPCR) polypeptide #54.	
XX		
XX	Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;	
KW	attention deficit disorder; anxiety; depression; bipolar disorder;	
KW	neurological disorder; Huntington's disease; dementia; obesity; anorexia;	
KW	metabolic disorder; Parkinson's disease; Tourette's syndrome; chromoblast;	
KW	type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;	
KW	cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;	
KW	viral infection; immunostimulant; neuroleptic; nootropic; tranquilizer;	
KW	antidepressant; anorectic; gene therapy.	
OS	Homo sapiens.	
XX		
PN	WO200162797-A2.	
XX		
PD	30-AUG-2001.	
XX		
PF	23-FEB-2001; 2001WO-US005676.	
XX		
PR	23-FEB-2000; 2000US-0184247P.	
XX		
PR	23-FEB-2000; 2000US-0184303P.	
XX		
PR	23-FEB-2000; 2000US-0184305P.	
XX		
PR	23-FEB-2000; 2000US-0184397P.	
XX		
PR	02-MAR-2000; 2000US-0186457P.	
XX		
PR	03-MAR-2000; 2000US-0186810P.	
XX		
PR	09-MAR-2000; 2000US-0188064P.	
XX		
PR	13-MAR-2000; 2000US-0194344P.	
XX		
PR	03-APR-2000; 2000US-0213861P.	
XX		
PR	23-JUN-2000; 2000US-0217369P.	
XX		
PR	11-JUL-2000; 2000US-0217370P.	
XX		
PR	14-JUL-2000; 2000US-0218337P.	
XX		
PR	20-JUL-2000; 2000US-0218492P.	
XX		
PA	(PHAA) PHARMACIA & UPJOHN CO.	
XX		
PI	Vogel G, Wood LS, Parodi LA, Lind P,	
XX		
DR	WPI; 2001-570628/64.	

DR N-PSDB; AAs42859.

XX New isolated nucleic acid encoding a new G-protein coupled receptor
 PT polypeptide for detecting receptor modulators that can treat mental
 PT disorders, such as schizophrenia, anxiety, depression, or obesity.

PS Claim 35; Page 91; 279pp; English.

XX Sequences AAU25554-AAU25616 represent human G-protein coupled receptor
 CC (GPCR) polypeptides of the invention. The proteins and their associated
 CC DNA sequences can be used to identify compounds which bind to GPCR
 CC polypeptides and in screening for compounds that modulate GPCR activity.
 CC By screening a human subject for the presence of mutations in GPCR DNA, a
 CC GPCR-related disorder or a genetic predisposition can be diagnosed. The
 CC sequences can also be used for treatment and prevention of mental
 CC disorders such as schizophrenia, attention deficit disorder, anxiety,
 CC depression, dementia and bipolar disorder, neurological disorders such as
 CC Huntington's disease, Parkinson's disease and Tourette's syndrome,
 CC metabolic disorders such as obesity, anorexia and type 2 diabetes,
 CC cardiovascular disorders such as thrombosis, myocardial infarction,
 CC cardiomyopathy and atherosclerosis, viral infections caused by HIV and
 CC cancers

SQ Sequence 333 AA;

Query Match 100.0%; Score 1733; DB 4; Length 333;

Best Local Similarity 100.0%; Pred. No. 4e-177; Mismatches 0; Gaps 0;

Matches 333; Conservative 0; Indels 0; Gaps 0;

QY 1 MMSGWFNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLANLHSAKASMT 60
 DB 1 MMSGWFNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLANLHSAKASMT 60

QY 61 MPDYVFVNNAAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVVALQIPNVSSLVAMYST 120
 DB 61 MPDYVFVNNAAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVVALQIPNVSSLVAMYST 120

QY 121 ALISLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALBCK 180
 DB 121 ALISLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALBCK 180

QY 121 ALISLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALBCK 180
 DB 121 ALISLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALBCK 180

QY 181 MONEAADATLVFIIGYVVPALATLYALVLSRVREDDPLDRDGRLEPSAHLIVATVC 240
 DB 181 MONEAADATLVFIIGYVVPALATLYALVLSRVREDDPLDRDGRLEPSAHLIVATVC 240

QY 241 TOFGMTPHYLILGHTVLIISRGKPVDAHYLGILHFPVDFSKLIAFSSSFVTPLIYRYNN 300
 DB 241 TOFGMTPHYLILGHTVLIISRGKPVDAHYLGILHFPVDFSKLIAFSSSFVTPLIYRYNN 300

QY 301 QSFPSKIQRLMKKLPCCGDRHCSPDHMGVOQVLA 333
 DB 301 QSFPSKIQRLMKKLPCCGDRHCSPDHMGVOQVLA 333

QY 301 QSFPSKIQRLMKKLPCCGDRHCSPDHMGVOQVLA 333
 DB 301 QSFPSKIQRLMKKLPCCGDRHCSPDHMGVOQVLA 333

RESULT 2

ABU04068 standard; protein; 333 AA.

ABU04068

11-OCT-2002 (first entry)

Human G protein coupled receptor hRUP28.

Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;

hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37.

Homo sapiens.

WO200242461-A2.

30-MAY-2002.

PF 26-NOV-2001; 2001WO-US044386.

XX 27-NOV-2001; 2000US-0253404P.

XX 12-DEC-2001; 2000US-0255366P.

XX 20-FEB-2001; 2001US-0270266P.

XX 20-FEB-2001; 2001US-0270266P.

XX 06-APR-2001; 2001US-0282032P.

XX 06-APR-2001; 2001US-0282358P.

XX 06-APR-2001; 2001US-0282358P.

XX 14-MAY-2001; 2001US-0290917P.

XX 31-JUL-2001; 2001US-0309208P.

XX (AREN-) ARENA PHARM INC.

XX Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;

XX WPI; 2002-566565/60.

XX N-PSDB; ABT04866.

XX Novel endogenous and non-endogenous versions of G protein-coupled

XX receptor useful for identification of candidate compounds as receptor

XX agonists or antagonists for use as therapeutic agents.

XX Claim 1; Page 55-56; 84pp; English.

SQ Sequence 333 AA;

Query Match 100.0%; Score 1733; DB 5; Length 333;

Best Local Similarity 100.0%; Pred. No. 4e-177; Mismatches 0; Gaps 0;

Matches 333; Conservative 0; Indels 0; Gaps 0;

QY 1 MMSGWFNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLANLHSAKASMT 60
 DB 1 MMSGWFNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLANLHSAKASMT 60

QY 61 MPDYVFVNNAAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVVALQIPNVSSLVAMYST 120
 DB 61 MPDYVFVNNAAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVVALQIPNVSSLVAMYST 120

QY 121 ALISLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALBCK 180
 DB 121 ALISLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALBCK 180

QY 121 ALISLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALBCK 180
 DB 121 ALISLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALBCK 180

QY 181 MONEAADATLVFIIGYVVPALATLYALVLSRVREDDPLDRDGRLEPSAHLIVATVC 240
 DB 181 MONEAADATLVFIIGYVVPALATLYALVLSRVREDDPLDRDGRLEPSAHLIVATVC 240

QY 241 TOFGMTPHYLILGHTVLIISRGKPVDAHYLGILHFPVDFSKLIAFSSSFVTPLIYRYNN 300
 DB 241 TOFGMTPHYLILGHTVLIISRGKPVDAHYLGILHFPVDFSKLIAFSSSFVTPLIYRYNN 300

QY 301 QSFPSKIQRLMKKLPCCGDRHCSPDHMGVOQVLA 333
 DB 301 QSFPSKIQRLMKKLPCCGDRHCSPDHMGVOQVLA 333

RESULT 3

AAU97734 standard; protein; 333 AA.

AAU97734

27-AUG-2002 (first entry)

Human TGR339 polypeptide.

KW Human: TGR339; receptor; G-protein coupled receptor; GPCR; TGR;
 KW TGR-associated disorder; signal transduction; renal failure; nephritis;
 KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;
 KW diabetes insipidus; hyperprolactinemia; thirst disturbance; appetite;
 KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;
 KW circadian rhythm.
 XX
 OS Homo sapiens.
 XX
 PN MO200242458-A2.
 XX
 PD 30-MAY-2002.
 XX
 PF 21-NOV-2001; 2001MO-US043404.
 XX
 PR 22-NOV-2000; 2000US-0252841P.
 PR 22-DEC-2000; 2000US-0257636P.
 PR 12-JAN-2001; 2001US-0261377P.
 PR 28-MAR-2001; 2001US-0279554P.
 PR 29-MAR-2001; 2001US-0280696P.
 XX
 PA (TULA-) TULARIK INC.
 XX
 PI Tien H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupta JS;
 XX
 DR WPI, 2002-463633/49.
 DR N-PSDB; ABR6286.
 XX
 PT New isolated G-protein couple receptor polypeptide, termed TGR, for
 PT diagnosis and treatment of diseases such as renal failure, nephritis,
 PT hypothyroidism, diabetes insipidus, and disturbances of thirst and sleep.
 XX
 PS Claim 33; Page 74; 98pp; English.
 XX
 XX The invention relates to a G-protein coupled receptor polypeptide (GPCR),
 CC termed TGR, and its associated nucleic acid. The sequences of the
 CC invention are useful for identifying a compound that modulates signal
 CC transduction and for identifying a mammal having a TGR-associated
 CC disorder. The proteins and nucleic acids are useful in diagnosis and
 CC treatment of diseases or conditions such as renal failure, nephritis,
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,
 CC diabetes insipidus, hyperprolactinemia and disturbances of thirst,
 CC sleep, temperature regulation, appetite, blood pressure or any other
 CC syndrome or disease associated with the hypothalamus. The sequences can
 CC be used in regulation of circadian rhythms, for use as genetic markers
 CC for the identification of mutations associated with diseases resulting
 CC from GPCR inactivation in particular cell types and for identification of
 CC modulators of GPCR signal transduction. This sequence represents the
 CC human TGR339 polypeptide
 CC
 XX
 SQ Sequence 333 AA;
 XX
 Query Match 100.0%; Score 1733; DB 5; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 MMSCSWNGTGLVEELPACQDLQGLSLSLGLVGVPGVGLCYNALLVTANLHRSKASMT 60
 DB 1 MMSCSWNGTGLVEELPACQDLQGLSLSLGLVGVPGVGLCYNALLVTANLHRSKASMT 60
 QY 61 MPVVYFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 120
 DB 61 MPVVYFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 120
 QY 121 ALSLDHYIERALPRTYMASVYNTTRHVCQVWGALLTSFSSLLFYICSHVSTRALBECAK 180
 DB 121 ALSLDHYIERALPRTYMASVYNTTRHVCQVWGALLTSFSSLLFYICSHVSTRALBECAK 180
 QY 181 MONEAADATLVFVIGVVPALATLVYALVLLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
 DB 181 MONEAADATLVFVIGVVPALATLVYALVLLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
 QY 241 TQGLMTPHYLLILGHTVITISRGKPVDAHYLGILHFKDPSKLLAPSSSPVTPLLVRYNM 300

DB 241 TQGLMTPHYLLILGHTVITISRGKPVDAHYLGILHFKDPSKLLAPSSSPVTPLLVRYNM 300
 QY 301 QSFPSLQRLMKKLPCCGRHPCSPDHMGVOQVLA 333
 DB 301 QSFPSLQRLMKKLPCCGRHPCSPDHMGVOQVLA 333
 XX
 RESULT 4
 ID ABR56303 standard; protein; 333 AA.
 XX
 AC ABR56303;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human G-protein coupled receptor 901.
 XX
 KW Human; anorectic; antidiabetic; antilipemic; hypothalamus;
 KW G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
 KW clobopobia; anorexia nervosa.
 XX
 OS Homo sapiens.
 XX
 PN WO2003030936-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-JP010250.
 XX
 PR 02-OCT-2001; 2001JP-00306872.
 XX
 PA (SUMU) SUMITOMO PHARM CO LTD.
 XX
 PI Suguru E, Tsuchida A, Yamanaka M, Tsiji M;
 XX
 DR WPI, 2003-354886/33.
 DR N-PSDB; ACC70841.
 XX
 PT Inhibitors of expression or activity of G-protein coupled receptor 901
 PT for treatment of lifestyle-related diseases and clobopobia.
 XX
 PS Claim 1; Page 64-65; 91pp; Japanese.
 XX
 XX The present invention relates to novel remedies for the treatment of
 CC diseases containing as an active component an inhibitor of the expression
 CC or activity of hypothalamus-expressed G-protein coupled receptor 901 and
 CC for treatment of clobopobia containing as an active component a
 CC potentiator of the expression or activity of G-protein coupled receptor
 CC 901. The diseases which can be treated include obesity, diabetes and
 CC hyperlipemia, and clobopobia (anorexia nervosa). The present human
 CC sequence is G-protein coupled receptor 901
 CC
 XX
 SQ Sequence 333 AA;
 XX
 Query Match 100.0%; Score 1733; DB 6; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 MMSCSWNGTGLVEELPACQDLQGLSLSLGLVGVPGVGLCYNALLVTANLHRSKASMT 60
 DB 1 MMSCSWNGTGLVEELPACQDLQGLSLSLGLVGVPGVGLCYNALLVTANLHRSKASMT 60
 QY 61 MPVVYFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 120
 DB 61 MPVVYFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 120
 QY 121 ALSLDHYIERALPRTYMASVYNTTRHVCQVWGALLTSFSSLLFYICSHVSTRALBECAK 180
 DB 121 ALSLDHYIERALPRTYMASVYNTTRHVCQVWGALLTSFSSLLFYICSHVSTRALBECAK 180
 QY 181 MONEAADATLVFVIGVVPALATLVYALVLLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
 DB 181 MONEAADATLVFVIGVVPALATLVYALVLLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240

DB 181 MONAADATLVFIGVVPALATLYALVLSRVRESDTPLDRTGRLPSAHLVAATVC 240
 QY 241 TOPGLWTPHYLLILGHTVLIISRGKVPDAHYLGILHFPVKDSKLLAFSSSPVTPPLLYRYNN 300
 DB 241 TOPGLWTPHYLLILGHTVLIISRGKVPDAHYLGILHFPVKDSKLLAFSSSPVTPPLLYRYNN 300
 QY 301 QSPFSKLORLMKKLPDGDHSCSPDHMGVOQVLA 333
 DB 301 QSPFSKLORLMKKLPDGDHSCSPDHMGVOQVLA 333

RESULT 5

ABU09900 standard; protein; 333 AA.
 ID ABU09900
 AC ABU09900;
 XX
 DT 11-AUG-2003 (first entry)
 XX

DE Human G-protein coupled receptor 23992.
 XX

KW Human; receptor; G-protein coupled receptor; gene therapy; pneumonia;
 KW immune disorder; Crohn's disease; Grave's disease; respiratory disorder;
 KW diphtheria; haematopoietic disorder; leukaemia; systemic sclerosis;
 KW prostatic disorder; benign prostate hypertrophy; tumour; breast disease;
 KW acute mastitis; Paget's disease; muscular disorder; rhabdomyosarcoma;
 KW neurological disorder; cerebral oedema; Parkinson's disease; atrophy;
 KW blood vessel disorder; atherosclerosis; testicular disease; syphilis;
 KW kidney disorder; cystic renal dysplasia; hyperthyroidism; cretinism; AIDS;
 KW cardiovascular disease; heart failure; pericarditis; pancreatic disease;
 KW pancreatitis; diabetes mellitus; thymus disease; thymic hypoplasia;
 KW Hodgkin disease; spleen disease; splenomegaly; Gaucher disease;
 KW liver disease; hepatic failure; alcoholic liver disease; T-cell disorder;
 KW systemic lupus erythematosus; B-cell disorder; Burkitt lymphoma;
 KW multiple myeloma; platelet disorder; thrombocytopenia;
 KW haemolytic-uraemic syndrome.
 XX

OS Homo sapiens.
 XX

PN US2003017539-A1.
 XX

PD 23-JAN-2003.
 XX

PF 07-JUN-2002; 2002US-00165844.
 XX

PR 02-JUN-1998; 98US-00088857.
 PR 02-SEP-1998; 98US-00145745.
 PR 21-JAN-1999; 99US-00234923.
 PR 02-JUN-1999; 99US-00324465.
 PR 28-JUN-1999; 99US-00340880.
 PR 26-AUG-1999; 99US-00383745.
 PR 16-DEC-1999; 99US-00464685.
 PR 18-DEC-2000; 2000US-00741783.
 XX

PA (MILL-) MILLENNIUM PHARM INC.
 XX

PI Glucksmann MA, Hodge MR, Hunter JV, Rudolph-Owen LA;
 PI SIllos-Santiago I, Welch NS;
 XX

DR MPI; 2003-401672/38.
 DR N-PSDB; ACA61176.
 XX

PT New nucleic acid molecule encoding a G-protein coupled receptor
 PT polypeptide, e.g. 2871, 14926, 17723 or 23992, useful for diagnosing
 PT and/or treating cancer or immune, respiratory, hematologic or
 PT cardiovascular disorders.
 XX

PS Claim 9; Fig 21; 14926; English.
 XX

CC The invention relates to an isolated G-protein coupled receptor nucleic
 CC acid molecule encoding the G-protein coupled receptors 2871, 14926, 17723
 CC and 23992. The nucleic acid molecule is useful in monitoring, diagnosing
 CC

CC and treating immune disorders (e.g. Crohn's disease and Grave's disease);
 CC respiratory disorders (e.g. pneumonia or diphtheria); haematopoietic
 CC disorders (e.g. leukaemia and systemic sclerosis); prostatic disorders
 CC (e.g. benign prostate hypertrophy and tumours); breast diseases (e.g.
 CC acute mastitis and Paget's disease); muscular disorders (e.g.
 CC rhabdomyosarcoma); neurological disorders (e.g. cerebral atrophy
 CC and syphilis); thyroid diseases (e.g. hyperthyroidism and cretinism);
 CC cardiovascular diseases (e.g. heart failure and glomerulonephritis);
 CC diseases (e.g. pancreatitis and diabetes mellitus); diseases involving
 CC the thymus (e.g. thymic hypoplasia and Hodgkin disease); diseases
 CC involving the liver (e.g. splenomegaly and Gaucher disease); diseases
 CC involving the liver (e.g. hepatic failure and alcoholic liver disease);
 CC cell disorders (e.g. systemic lupus erythematosus and AIDS); B-cell
 CC disorder (e.g. Burkitt lymphoma and multiple myeloma); platelet disorders
 CC (e.g. thrombocytopenia and haemolytic-uraemic syndrome). The nucleic
 CC acids may also be used in chromosome mapping, tissue typing,
 CC pharmacogenomics and forensic biology, and as surrogate markers. The
 CC present sequence represents the amino acid sequence of the human G-
 CC protein coupled receptor 23992
 XX

Sequence 333 AA;

Query Match 100.0%; Score 1733; DB 6; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSGWFNGTGVLELPAQODLQGLSLSLGLVGVGPVGLCYNNLLVLANHSKASMT 60
 DB 1 MWSGWFNGTGVLELPAQODLQGLSLSLGLVGVGPVGLCYNNLLVLANHSKASMT 60
 QY 61 MPDYFVMAVAGVLSBALPVHLGPPSSRMALNSYGEVHTALQIPNVSSLVAMYST 120
 DB 61 MPDYFVMAVAGVLSBALPVHLGPPSSRMALNSYGEVHTALQIPNVSSLVAMYST 120
 QY 121 ALLSDHYIERALPRTYMASVNTTRHVCGFPMGALLTFSSLLFYICSHVSTRALMECK 180
 DB 121 ALLSDHYIERALPRTYMASVNTTRHVCGFPMGALLTFSSLLFYICSHVSTRALMECK 180
 QY 181 MONAADATLVFIGVVPALATLYALVLSRVRESDTPLDRTGRLPSAHLVAATVC 240
 DB 181 MONAADATLVFIGVVPALATLYALVLSRVRESDTPLDRTGRLPSAHLVAATVC 240
 QY 241 TOPGLWTPHYLLILGHTVLIISRGKVPDAHYLGILHFPVKDSKLLAFSSSPVTPPLLYRYNN 300
 DB 241 TOPGLWTPHYLLILGHTVLIISRGKVPDAHYLGILHFPVKDSKLLAFSSSPVTPPLLYRYNN 300
 QY 301 QSPFSKLORLMKKLPDGDHSCSPDHMGVOQVLA 333
 DB 301 QSPFSKLORLMKKLPDGDHSCSPDHMGVOQVLA 333

RESULT 6

ADC86563
 ID ADC86563 standard; protein; 333 AA.
 XX

AC ADC86563;
 XX

DT 01-JAN-2004 (first entry)
 XX

DE Human GPCR protein SEQ ID NO:1016.
 XX

KW human; GPCR; guanosine triphosphate-binding protein, coupled receptor;
 KW gene therapy.
 XX

OS Homo sapiens.
 XX

PN EP1270724-A2.
 XX

PD 02-JAN-2003.
 XX

PF 18-JUN-2002; 2002EP-00013517.
 XX
 XX 18-JUN-2001; 2001JP-00246789.
 XX
 PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATORIO.
 XX
 PI Suwa M, Asai K, Akiyama Y, Aburatani H;
 DR WPI, 2003-315783/31.
 DR N-PSDB; ADC86562.
 XX
 PT New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 XX
 PS Claim 2; SEQ ID NO 1016; 28pp; English.
 XX
 CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in ADC8549-ADC87617 represent GPCR's of the invention.
 XX
 SQ Sequence 333 AA.
 XX
 Query Match 100.0%; Score 1733; DB 7; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 MMSCSWNGTGVLELPAQCDLQGLSLSLGLGVGVPVGLCYNALVLANLHRSKASMT 60
 DB 1 MMSCSWNGTGVLELPAQCDLQGLSLSLGLGVGVPVGLCYNALVLANLHRSKASMT 60
 QY 61 MPVYEFNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPNVSSLVAMYST 120
 DB 61 MPVYEFNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPNVSSLVAMYST 120
 QY 121 ALLSLDHYIRALPRTYMASVYTRHVGCFVWGALLTSPSSLLFYICSHVSTRALECAK 180
 DB 121 ALLSLDHYIRALPRTYMASVYTRHVGCFVWGALLTSPSSLLFYICSHVSTRALECAK 180
 QY 181 MONEAADAATLVPIGYVVPALATLYALVLSRVRRETPIDRDTGLEPSAHLVATVC 240
 DB 181 MONEAADAATLVPIGYVVPALATLYALVLSRVRRETPIDRDTGLEPSAHLVATVC 240
 QY 241 TORGLMTPHYLILGHVTVIISRGKPVDAHVLGLHFKDSSKILAFSSSVTPLLVRYNM 300
 DB 241 TORGLMTPHYLILGHVTVIISRGKPVDAHVLGLHFKDSSKILAFSSSVTPLLVRYNM 300
 QY 301 QSFPSKLGRLMKKLPCCDRHCSPDHMGVQVLA 333
 DB 301 QSFPSKLGRLMKKLPCCDRHCSPDHMGVQVLA 333
 XX
 RESULT 7
 ADO29606
 ID ADO29606 standard; protein; 333 AA.
 AC ADO29606;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human GPCR PGR8, SEQ ID NO:708.
 XX
 KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;

KM ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KM skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KM thymus disorder; thyroid disorder; antiparkinsonian; antineuritic;
 KM cytoacetic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
 KM CNS; central nervous system; respiratory; antidiabetic; antidiabetic;
 KM virucide; hepatotropic; antibacterial; antianemic; antiseborrheic;
 KM dermatological; antitumor; antihypertensive; anorectic;
 KM immunosuppressive; nephrotoxic; gene therapy; GPCR modulator; human;
 KM receptor.
 XX
 OS Homo sapiens.
 XX
 PN W02004040000-A2.
 XX
 PD 13-MAY-2004.
 XX
 PE 09-SEP-2003; 2003WO-US028226.
 XX
 PR 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX
 PA (PRIM-) PRIMAL INC.
 XX
 PI Galanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
 PI Medisen L, McIlwain KL, Pavlova MN, Vasiliadis D, Zeng H;
 DR WPI, 2004-390329/36.
 DR N-PSDB; ADO30022.
 XX
 PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX
 PS Claim 151; SEQ ID NO 708; 542pp; English.
 XX
 CC The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders (e.g., autoimmune disorders or
 CC anemias or leukemias); immune disorders (e.g., osteoarthritis, rheumatoid
 CC AIDS); bone and joint disorders (e.g., osteoporosis); metabolic or nutritive disorders (e.g.,
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast; ovary;
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancer). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 333 AA;
 XX
 Query Match 100.0%; Score 1733; DB 8; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWENGSTGLVEELPACODLQIGLSLSLGLVGVPGVGLCYNNALLVIANLHKSASMT 60
 DB 1 MMSCSWENGSTGLVEELPACODLQIGLSLSLGLVGVPGVGLCYNNALLVIANLHKSASMT 60
 QY 61 MPDYVFVMAVAGLVLSALAPVHLIGPSSRMALMSVCGEVHVAIQIPFNVSLLVAMYST 120
 DB 61 MPDYVFVMAVAGLVLSALAPVHLIGPSSRMALMSVCGEVHVAIQIPFNVSLLVAMYST 120
 QY 121 ALLSDHYIERALPRTYMAVYNTRHVCGFWGCGALLTSSSLLFYICSHVSTRALBECAK 180
 DB 121 ALLSDHYIERALPRTYMAVYNTRHVCGFWGCGALLTSSSLLFYICSHVSTRALBECAK 180
 QY 181 MONEADATLVFVIGVVPALATLYALVLSRVRREDTLPDRDGRLEPSAHRLLVATVC 240
 DB 181 MONEADATLVFVIGVVPALATLYALVLSRVRREDTLPDRDGRLEPSAHRLLVATVC 240
 QY 241 TQFGLMTPHYLLILGHTVILISRGKPDVAHYLGHLHFVKDPSKLLAFSSSFVTPLLYRYMN 300
 DB 241 TQFGLMTPHYLLILGHTVILISRGKPDVAHYLGHLHFVKDPSKLLAFSSSFVTPLLYRYMN 300
 QY 301 QSFPSKIQRLMKKLPCCDRHCSPDHMGVQVLA 333
 DB 301 QSFPSKIQRLMKKLPCCDRHCSPDHMGVQVLA 333

RESULT 8
 AAM25659
 ID AAM25659 standard; protein; 370 AA.
 AC AAM25659;
 XX
 DT 16-OCT-2001 (first entry)
 DE
 XX
 DE Human protein sequence SEQ ID NO:1174.
 XX
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW anti-infective; antineoplastic; antiarrhythmic; immunosuppressive;
 KW antibacterial; endocrine; cardiant; central nervous system; vitruide;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KW antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;
 KW dermatological; antiallergic; antiasthmatic; antiparkinsonian; infection;
 KW neuroprotective; antidiabetic; nootropic; antidiabetic; cytostatic;
 KW immunostimulant; gene therapy; antitumor; antiparkinsonian; infection;
 KW immunomodulatory; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PE 22-DEC-2000; 2000WO-US035017.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-457603/49.
 XX
 DR N-PSDB; AAH99600.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 treatment and diagnosis of e.g. cancer, cancer, ulcers and HIV infection.

XX
 PS Claim 20; Page 244; 1217pp; English.
 XX
 CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: anti-inflammatory; antineoplastic;
 CC antiarrhythmic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
 CC antitumor; osteopathic; dermatological; antiallergic; antiasthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidiabetic; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders
 CC
 SQ Sequence 370 AA:
 Query Match 100.0%; Score 1733; DB 4; Length 370;
 Best Local Similarity 100.0%; Pred. No. 4, 6e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MMSCSWENGSTGLVEELPACODLQIGLSLSLGLVGVPGVGLCYNNALLVIANLHKSASMT 60
 DB 38 MMSCSWENGSTGLVEELPACODLQIGLSLSLGLVGVPGVGLCYNNALLVIANLHKSASMT 97
 QY 61 MPDYVFVMAVAGLVLSALAPVHLIGPSSRMALMSVCGEVHVAIQIPFNVSLLVAMYST 120
 DB 98 MPDYVFVMAVAGLVLSALAPVHLIGPSSRMALMSVCGEVHVAIQIPFNVSLLVAMYST 157
 QY 121 ALLSDHYIERALPRTYMAVYNTRHVCGFWGCGALLTSSSLLFYICSHVSTRALBECAK 180
 DB 158 ALLSDHYIERALPRTYMAVYNTRHVCGFWGCGALLTSSSLLFYICSHVSTRALBECAK 217
 QY 181 MONEADATLVFVIGVVPALATLYALVLSRVRREDTLPDRDGRLEPSAHRLLVATVC 240
 DB 218 MONEADATLVFVIGVVPALATLYALVLSRVRREDTLPDRDGRLEPSAHRLLVATVC 277
 QY 241 TQFGLMTPHYLLILGHTVILISRGKPDVAHYLGHLHFVKDPSKLLAFSSSFVTPLLYRYMN 300
 DB 278 TQFGLMTPHYLLILGHTVILISRGKPDVAHYLGHLHFVKDPSKLLAFSSSFVTPLLYRYMN 337
 QY 301 QSFPSKIQRLMKKLPCCDRHCSPDHMGVQVLA 333
 DB 338 QSFPSKIQRLMKKLPCCDRHCSPDHMGVQVLA 370

RESULT 9
 ABR56305
 ID ABR56305 standard; protein; 695 AA.
 AC ABR56305;
 XX
 DT 20-NOV-2003 (first entry)
 DE
 XX
 DE pc90HISGalpa12 protein.
 XX
 KW Human; anorectic; antidiabetic; antilipemic; hypothalamus;
 KW G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
 KW cibophobia; anorexia nervosa.
 XX
 OS Unidentified.
 XX
 PN WO2003030936-A1.

XX 17-APR-2003.
PD 02-OCT-2002; 2002MO-JP010250.
XX PF 02-OCT-2001; 2001JP-00306872.
XX PR 02-OCT-2001; 2001JP-00306872.
XX (SUMU) SUMITOMO PHARM CO LTD.
XX PA Suguru E, Tsuchida A, Yamanaka M, Taiji M;
XX WPI; 2003-354886/33.
XX DR N-PSDB; ACC70860.
XX PT Inhibitors of expression or activity of G-protein coupled receptor 901
XX for treatment of lifestyle-related diseases and cibophobia.
XX PS Disclosure; Page 79-81, 91pp; Japanese.
XX CC The present invention relates to novel remedies for the treatment of
XX diseases containing as an active component an inhibitor of the expression
XX or activity of hypothalamus-expressed G-protein coupled receptor 901 and
XX CC for treatment of cibophobia containing as an active component a
XX CC potentiator of the expression or activity of G-protein coupled receptor
XX 901. The diseases which can be treated include obesity, diabetes and
XX CC hyperlipemia, and cibophobia (anorexia nervosa). The present sequence
XX was used to illustrate the invention
SQ Sequence 695 AA;

Query Match 100.0%; Score 1733; DB 6; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.1e-176;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVLANLHKSASMT 60
DB 1 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVLANLHKSASMT 60
QY 61 MPVYFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGSVHVVALQIPFNVSLSLVAMYST 120
DB 61 MPVYFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGSVHVVALQIPFNVSLSLVAMYST 120
QY 121 ALSLDHYIERALPRTYMASVNTTRHYCGFWGALLTSFSSLLFYICSHVSTRALBCKA 180
DB 121 ALSLDHYIERALPRTYMASVNTTRHYCGFWGALLTSFSSLLFYICSHVSTRALBCKA 180
QY 181 MONEAADATLVFVIGVVPALATLVALLSRVRREDTPLDRDTGRLEPSAHLVAATVC 240
DB 181 MONEAADATLVFVIGVVPALATLVALLSRVRREDTPLDRDTGRLEPSAHLVAATVC 240
QY 241 TORGLMTPHYLLILGHTVILSRGKPDVAHYLGILHFKDPSKLIAPSSSVFTPLLYRYNN 300
DB 241 TORGLMTPHYLLILGHTVILSRGKPDVAHYLGILHFKDPSKLIAPSSSVFTPLLYRYNN 300
QY 301 QSFPSKLGRLMKKLPBGDRHCSPDHMGVQOYLA 333
DB 301 QSFPSKLGRLMKKLPBGDRHCSPDHMGVQOYLA 333

RESULT 10
ABR56304
ID ABR56304 standard; protein; 714 AA.
XX ABR56304;
AC ABR56304;
XX 20-NOV-2003 (first entry)
XX pc901HISga1pha16 protein.
XX Human; anorectic; antidiabetic; antilipemic; hypothalamus;
XX G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
XX cibophobia; anorexia nervosa.
XX

OS Unidentified.
XX XX
XX W02003030936-A1.
XX
XX 17-APR-2003.
XX PD 02-OCT-2002; 2002MO-JP010250.
XX PF 02-OCT-2001; 2001JP-00306872.
XX PR 02-OCT-2001; 2001JP-00306872.
XX (SUMU) SUMITOMO PHARM CO LTD.
XX PA Suguru E, Tsuchida A, Yamanaka M, Taiji M;
XX WPI; 2003-354886/33.
XX DR N-PSDB; ACC70859.
XX PT Inhibitors of expression or activity of G-protein coupled receptor 901
XX for treatment of lifestyle-related diseases and cibophobia.
XX PS Disclosure; Page 73-75, 91pp; Japanese.
XX CC The present invention relates to novel remedies for the treatment of
XX diseases containing as an active component an inhibitor of the expression
XX or activity of hypothalamus-expressed G-protein coupled receptor 901 and
XX CC for treatment of cibophobia containing as an active component a
XX CC potentiator of the expression or activity of G-protein coupled receptor
XX 901. The diseases which can be treated include obesity, diabetes and
XX CC hyperlipaemia, and cibophobia (anorexia nervosa). The present sequence
XX was used to illustrate the invention
SQ Sequence 714 AA;

Query Match 100.0%; Score 1733; DB 6; Length 714;
Best Local Similarity 100.0%; Pred. No. 1.1e-176;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVLANLHKSASMT 60
DB 1 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVLANLHKSASMT 60
QY 61 MPVYFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGSVHVVALQIPFNVSLSLVAMYST 120
DB 61 MPVYFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGSVHVVALQIPFNVSLSLVAMYST 120
QY 121 ALSLDHYIERALPRTYMASVNTTRHYCGFWGALLTSFSSLLFYICSHVSTRALBCKA 180
DB 121 ALSLDHYIERALPRTYMASVNTTRHYCGFWGALLTSFSSLLFYICSHVSTRALBCKA 180
QY 181 MONEAADATLVFVIGVVPALATLVALLSRVRREDTPLDRDTGRLEPSAHLVAATVC 240
DB 181 MONEAADATLVFVIGVVPALATLVALLSRVRREDTPLDRDTGRLEPSAHLVAATVC 240
QY 241 TORGLMTPHYLLILGHTVILSRGKPDVAHYLGILHFKDPSKLIAPSSSVFTPLLYRYNN 300
DB 241 TORGLMTPHYLLILGHTVILSRGKPDVAHYLGILHFKDPSKLIAPSSSVFTPLLYRYNN 300
QY 301 QSFPSKLGRLMKKLPBGDRHCSPDHMGVQOYLA 333
DB 301 QSFPSKLGRLMKKLPBGDRHCSPDHMGVQOYLA 333

RESULT 11
ABR56306
ID ABR56306 standard; protein; 735 AA.
XX ABR56306;
AC ABR56306;
XX 20-NOV-2003 (first entry)
XX pc901HISga1pha2 protein.
XX Human; anorectic; antidiabetic; antilipemic; hypothalamus;
XX

KM G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
KW cibophobia; anorexia nervosa.
XX
OS Unidentified.
XX
FN WO2003030936-A1.
XX
PD 17-APR-2003.
XX
PF 02-OCT-2002; 2002WO-JP010250.
XX
PR 02-OCT-2001; 2001JP-00306872.
XX
PA (SUMU) SUMITOMO PHARM CO LTD.
XX
PI Suguru E, Teuchida A, Yamanaoka M, Taiji M;
XX
DR WPI; 2003-354886/33.
XX
DR N-PSDB; ACC70861.
XX
PT Inhibitors of expression or activity of G-protein coupled receptor 901
XX for treatment of lifestyle-related diseases and cibophobia.
XX
PS Disclosure; Page 85-87; 91pp; Japanese.
XX
CC The present invention relates to novel remedies for the treatment of
CC diseases containing an active component an inhibitor of the expression
CC or activity of hypothalamus-expressed G-protein coupled receptor 901 and
CC for treatment of cibophobia containing an active component a
CC potentiator of the expression or activity of G-protein coupled receptor
CC 901. The diseases which can be treated include obesity, diabetes and
CC hyperlipaemia, and cibophobia (anorexia nervosa). The present sequence
CC was used to illustrate the invention
XX
SQ Sequence 735 AA;
Query Match 100.0%; Score 1733; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.2e-176;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWSGSMFNGTGLVEELPACODLQGLSLSLGLVGVVGVGVCYNALLVLANLHSAKSMT 60
DB 1 MWSGSMFNGTGLVEELPACODLQGLSLSLGLVGVVGVGVCYNALLVLANLHSAKSMT 60
QY 61 MPDYTFVNMAVAGLVLSALAPVHLLGPSSRNALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYTFVNMAVAGLVLSALAPVHLLGPSSRNALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSLDHYTERALPRTYMASVYNTTRHVCFFWGGALLTFSSLLFTYICSHVSTRALAECAK 180
DB 121 ALLSLDHYTERALPRTYMASVYNTTRHVCFFWGGALLTFSSLLFTYICSHVSTRALAECAK 180
QY 181 MONAERADATLVFVIGVVPALATLVALLVLSRVRREDTFLDSDTGRLEBSAHLVATVC 240
DB 181 MONAERADATLVFVIGVVPALATLVALLVLSRVRREDTFLDSDTGRLEBSAHLVATVC 240
QY 241 TQFGLMTPHYLILGHTVILISGKPYDAHYLGILHFVXDFSKLAFSSFTVPLLYRYMN 300
DB 241 TQFGLMTPHYLILGHTVILISGKPYDAHYLGILHFVXDFSKLAFSSFTVPLLYRYMN 300
QY 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333
RESULT 12
AAV57285
XX AAV57285 standard; protein; 333 AA.
XX
AC AAV57285;
XX
DT 05-JUN-2000 (first entry)
XX

DE Human GPCR protein (HGPR) sequence (clone ID 2214673).
XX
KW Human; G protein coupled protein receptor; HGPR; cell proliferation;
KW neurological; immune disorder; cytostatic; anti-arteriosclerotic;
KW anti-atherosclerotic; hepatotropic; antiinflammatory; virocidic; leukemia;
KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;
KW anti-Alzheimer's; anti-Parkinsonian; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200015793-A2.
XX
PD 23-MAR-2000.
XX
PF 17-SEP-1999; 99WO-US020958.
XX
PR 17-SEP-1998; 98US-00156513.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Bandman O, Lai P, Tang YT, Corley NC, Guegler KJ, Gorgone GA,
XX PI Baughn MR;
XX
DR WPI; 2000-271432/23.
XX
DR N-PSDB; AA290523.
XX
PT Human G protein coupled protein receptor peptides useful for the
XX prevention, diagnosis and treatment of cell proliferative, neurological
XX and immune disorders.
XX
PS Claim 1; Page 68-69; 71pp; English.
XX
CC The invention provides human G protein coupled protein receptor (HGPR)
CC polypeptides and polynucleotides encoding them. The polypeptides can be
CC produced by standard recombinant methodology. The polynucleotides and
CC polypeptides may be used in the prevention, treatment and diagnosis of
CC diseases associated with their inappropriate expression. Diseases that
CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
CC and Parkinson's disease). The anti-HGPR antibodies may be used as
CC diagnostic agents for detecting the presence of HGPR polypeptides in
CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences
XX AAV57283-288 represent the HGPR polypeptides
XX
SQ Sequence 333 AA;
Query Match 99.6%; Score 1726; DB 3; Length 333;
Best Local Similarity 99.7%; Pred. No. 2.2e-176;
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MWSGSMFNGTGLVEELPACODLQGLSLSLGLVGVVGVGVCYNALLVLANLHSAKSMT 60
DB 1 MWSGSMFNGTGLVEELPACODLQGLSLSLGLVGVVGVGVCYNALLVLANLHSAKSMT 60
QY 61 MPDYTFVNMAVAGLVLSALAPVHLLGPSSRNALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYTFVNMAVAGLVLSALAPVHLLGPSSRNALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSLDHYTERALPRTYMASVYNTTRHVCFFWGGALLTFSSLLFTYICSHVSTRALAECAK 180
DB 121 ALLSLDHYTERALPRTYMASVYNTTRHVCFFWGGALLTFSSLLFTYICSHVSTRALAECAK 180
QY 181 MONAERADATLVFVIGVVPALATLVALLVLSRVRREDTFLDSDTGRLEBSAHLVATVC 240
DB 181 MONAERADATLVFVIGVVPALATLVALLVLSRVRREDTFLDSDTGRLEBSAHLVATVC 240
QY 241 TQFGLMTPHYLILGHTVILISGKPYDAHYLGILHFVXDFSKLAFSSFTVPLLYRYMN 300
DB 241 TQFGLMTPHYLILGHTVILISGKPYDAHYLGILHFVXDFSKLAFSSFTVPLLYRYMN 300
QY 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333

Db 301 QSFPSKQLRLMKKLLPCGDRHCSPDHMGVQVLA 333

RESULT 13
AAU25584
ID AAU25584 standard; protein; 333 AA.
XX
XX AAU25584;
DT 18-DEC-2001 (first entry)
XX
XX Human G Protein-Coupled Receptor (GPCR) polypeptide #31.
XX
XX Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;
XX attention deficit disorder; anxiety; depression; bipolar disorder;
XX neurological disorder; Huntington's disease; dementia; obesity; anorexia;
XX metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;
XX type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;
XX cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;
XX viral infection; immunostimulant; neuroleptic; nootropic; tranquiliser;
XX antidepressant; anorectic; gene therapy.
XX
XX Homo sapiens.
XX
XX MO200162797-A2.
PD 30-AUG-2001.
XX
XX 23-FEB-2001; 2001WO-US005676.
XX
XX 23-FEB-2000; 2000US-0184247P.
PR 23-FEB-2000; 2000US-0184303P.
PR 23-FEB-2000; 2000US-0184304P.
PR 23-FEB-2000; 2000US-0184305P.
PR 23-FEB-2000; 2000US-0184397P.
PR 02-MAR-2000; 2000US-0186457P.
PR 03-MAR-2000; 2000US-0186810P.
PR 09-MAR-2000; 2000US-0188064P.
PR 13-MAR-2000; 2000US-0188880P.
PR 03-APR-2000; 2000US-0194344P.
PR 23-JUN-2000; 2000US-0213861P.
PR 11-JUL-2000; 2000US-0217369P.
PR 11-JUL-2000; 2000US-0217370P.
PR 14-JUL-2000; 2000US-0218337P.
PR 20-JUL-2000; 2000US-0218492P.
XX
XX (PHAA) PHARMACIA & UPJOHN CO.
XX
XX
PI Vogel1 G, Wood LS, Parodi LA, Lind P;
XX
XX WPI; 2001-570628/64.
DR N-PSDB; AAS42836.
XX
XX
PT New isolated nucleic acid encoding a new G-protein coupled receptor
PT polypeptide for detecting receptor modulators that can treat mental
PT disorders, such as schizophrenia, anxiety, depression, or obesity.
XX
XX
PS Claim 35; Page 81; 279pp; English.

CC cancers
XX
XX Sequence 333 AA;
SQ
Query Match 99.6%; Score 1726; DB 4; Length 333;
Best Local Similarity 99.7%; Pred. No. 2.2e-176;
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MMSCSWNGTGVVEELPACODLQLGLSLSLGLVGVGVGLCYNALLVLANHRSASMT 60
DB 1 MMSCSWNGTGVVEELPACODLQLGLSLSLGLVGVGVGLCYNALLVLANHRSASMT 60
QY 61 MEDVYEVMAVAGLVLSALAPVHLGLPSSRRALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MEDVYEVMAVAGLVLSALAPVHLGLPSSRRALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSSSLLFTYICSHVSTRALECAK 180
DB 121 ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSSSLLFTYICSHVSTRALECAK 180
QY 181 MQNAEAADATLVPIGVVVALATLVLYLSRVREDTFLDRDTEGLSESAHRLVAVTC 240
DB 181 MQNAEAADATLVPIGVVVALATLVLYLSRVREDTFLDRDTEGLSESAHRLVAVTC 240
QY 241 TOPGLWTPHYLLIGHTVLIISRGKPYDAHYGLHFFVDKFSKLLAFSSFTPELYRYNM 300
DB 241 TOPGLWTPHYLLIGHTVLIISRGKPYDAHYGLHFFVDKFSKLLAFSSFTPELYRYNM 300
QY 301 QSFPSKQLRLMKKLLPCGDRHCSPDHMGVQVLA 333
DB 301 QSFPSKQLRLMKKLLPCGDRHCSPDHMGVQVLA 333
RESULT 14
ABU54867
ID ABU54867 standard; protein; 333 AA.
XX
XX ABU54867;
XX
XX 14-MAR-2003 (first entry)
DT
XX
XX Human G-protein coupled receptor #3, Incyte clone 2214673CD1.
DB
XX
XX Human; G-protein coupled receptor; GPCR; neoplastic disorder;
XX neurological disorder; immune disorder; cytostatic; pancreatic cancer;
XX follicular carcinoma of the thyroid; leiomyoma of the uterus; epilepsy;
XX interstitial nephritis; cancer.
XX
XX
OS Homo sapiens.
XX
XX
PN US2002106655-A1.
XX
XX 08-AUG-2002.
PD
XX
XX 28-JUN-2001; 2001US-00895686.
PF
XX
XX 17-SEP-1998; 98US-00156513.
PR
XX
XX (BAND/) BANDMAN O.
PA (LALP/) LAL P G.
PA (TANG/) TANG Y T.
PA (BAUG/) BAUGHN M R.
PI Bandman O, Lal PG, Tang YT, Baughn MR;
XX
XX WPI; 2002-697866/75.
DR N-PSDB; ABX73053.
XX
XX
PT New cDNAs encoding G protein coupled receptors are useful for the
PT diagnosis, prognosis, treatment and evaluation of therapies for
PT neoplastic, neurological and immune disorders.
XX
XX
PS Claim 13; Fig 2; 61pp; English.

The invention relates to an isolated cDNA encoding G-protein coupled receptor (GPCR) appearing as ABUS4865-ABUS4870. Also included are fragments of the cDNAs, species variants having at least 75% identity to the cDNAs, vectors comprising the cDNAs, a host cells comprising the above vectors, producing a protein (comprising culturing the above host cell under expression conditions and recovering the protein), using a cDNA to detect expression of a nucleic acid in a sample or to screen for compounds or molecules which bind to the cDNAs, using the GPCR protein to screen compounds or molecules for ligands, using a GPCR protein to prepare and purify antibodies, an anti-GPCR antibody and using the antibody to detect expression of a GPCR protein in a sample and its diagnostic of cancer. The invention is useful for the diagnosis, prognosis, treatment and evaluation of therapies for neoplastic, neurological and immune disorders, particularly follicular carcinoma of the thyroid, leiomyoma of the uterus, pancreatic cancer, epilepsy, interstitial nephritis and immune responses as a complication of cancer. The present sequence is a human GPCR protein of the invention

Sequence 333 AA:

OY 301 OSFPSKLORLMKXLP CGDRHCS PDDHMGVQOYLA 333
Db 301 OSFPSKLORLMKXLP CGDRHCS PDDHMGVQOYLA 333

Search completed: December 14, 2004, 20:23:05
Job time : 56.3965 sec8

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:25:14 ; Search time 15.535 Seconds
(without alignments)
1421.692 Million cell updates/sec

Title: US-09-995-225B-2

Perfect score: 1733
Sequence: 1 MMSCSWFMNGTGLVEELPACQ.....LPCGDRHCPDHMGVQCVLA 333

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 aeqa, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB aeq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodaca/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodaca/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodaca/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodaca/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodaca/1/1aa/6C.COMB.pep: *
6: /cgn2_6/ptodaca/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1711	98.7	333	4 US-09-409-778-3	Sequence 3, Appl1
2	271	15.6	375	2 US-08-480-994-38	Sequence 38, Appl1
3	271	15.6	375	2 US-08-616-844-38	Sequence 38, Appl1
4	271	15.6	375	2 US-08-559-654-38	Sequence 38, Appl1
5	271	15.6	375	2 US-08-485-573-38	Sequence 38, Appl1
6	271	15.6	375	3 US-08-944-868A-38	Sequence 38, Appl1
7	271	15.6	375	3 US-08-944-423A-38	Sequence 38, Appl1
8	271	15.6	375	3 US-08-925-743-38	Sequence 38, Appl1
9	271	15.6	375	3 US-08-944-496-38	Sequence 38, Appl1
10	271	15.6	375	3 US-08-925-767-38	Sequence 38, Appl1
11	269	15.5	375	4 US-09-170-496D-194	Sequence 194, App
12	269	15.5	375	4 US-09-170-496D-202	Sequence 202, App
13	267	15.4	375	4 US-09-711-889-2	Sequence 2, Appl1
14	267	15.4	375	4 US-09-170-068-2	Sequence 2, Appl1
15	267	15.4	375	4 US-09-170-496D-56	Sequence 56, Appl1
16	267	15.4	375	4 US-09-170-496D-70	Sequence 70, Appl1
17	200	11.5	355	1 US-07-759-568-1	Sequence 8, Appl1
18	200	11.5	355	1 US-08-446-669-8	Sequence 8, Appl1
19	200	11.5	355	2 US-08-390-000A-5	Sequence 8, Appl1
20	200	11.5	355	2 US-08-446-669-8	Sequence 8, Appl1
21	200	11.5	355	4 US-09-625-573-8	Sequence 8, Appl1
22	200	11.5	355	5 PCT-US95-00476-8	Sequence 8, Appl1
23	200	11.5	360	1 US-08-202-056-7	Sequence 7, Appl1
24	200	11.5	360	4 US-09-409-778-4	Sequence 4, Appl1
25	196.5	11.3	312	1 US-08-118-270-38	Sequence 38, Appl1
26	196.5	11.3	312	5 PCT-US93-08528-38	Sequence 38, Appl1
27	193.5	11.2	350	1 US-08-202-056-1	Sequence 1, Appl1

28	193.5	11.2	350	1 US-08-076-093A-2	Sequence 2, Appl1
29	193.5	11.2	350	1 US-08-450-393A-7	Sequence 7, Appl1
30	193.5	11.2	350	1 US-08-410-453A-1	Sequence 1, Appl1
31	193.5	11.2	350	1 US-08-701-265-2	Sequence 2, Appl1
32	193.5	11.2	350	1 US-08-410-454A-1	Sequence 1, Appl1
33	193.5	11.2	350	2 US-08-284-586-2	Sequence 2, Appl1
34	193.5	11.2	350	2 US-08-410-456A-1	Sequence 1, Appl1
35	193.5	11.2	350	2 US-08-805-478-2	Sequence 2, Appl1
36	193.5	11.2	350	2 US-08-802-627A-2	Sequence 2, Appl1
37	193.5	11.2	350	2 US-08-801-238-2	Sequence 2, Appl1
38	193.5	11.2	350	2 US-08-801-228-2	Sequence 2, Appl1
39	193.5	11.2	350	3 US-09-104-296-2	Sequence 7, Appl1
40	193.5	11.2	350	3 US-08-446-669-7	Sequence 7, Appl1
41	193.5	11.2	350	4 US-09-625-573-7	Sequence 7, Appl1
42	193.5	11.2	350	4 PCT-US95-00476-7	Sequence 7, Appl1
43	191.5	11.1	355	4 US-09-503-219B-8	Sequence 8, Appl1
44	188.5	10.9	365	4 US-10-039-659A-10	Sequence 10, Appl1
45	184.5	10.6	374	3 US-09-045-583-48	Sequence 48, Appl1

ALIGNMENTS

RESULT 1
US-09-409-778-3

Sequence 3, Application US/09409778

Patent No. 6472173

GENERAL INFORMATION:

APPLICANT: Ford, John

APPLICANT: Yeung, George

TITLE OF INVENTION: A NOVEL CHEMOKINE RECEPTOR OBTAINED FROM

TITLE OF INVENTION: A CDNA LIBRARY OF FETAL LIVER-SPLEEN

FILE REFERENCE: 20411-742CON2 (now 28110/36057B)

CURRENT APPLICATION NUMBER: US/09/409,778

PRIOR APPLICATION NUMBER: PCT/US99/12829

PRIOR FILING DATE: 1999-06-29

PRIOR APPLICATION NUMBER: US 09/236,166

PRIOR FILING DATE: 1999-01-22

PRIOR APPLICATION NUMBER: US 09/106,800

PRIOR FILING DATE: 1998-06-26

NUMBER OF SEQ IDS NOS: 25

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 3

LENGTH: 333

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1...(333))

OTHER INFORMATION: Xaa = Any Amino Acid

US-09-409-778-3

Query Match

Best Local Similarity 98.7%; Score 1711; DB 4; Length 333;

Matchea 330; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	MMSCSWFMNGTGLVEELPACQDQLGLSLSLGLVGVPGVLCYNMLLVLANIHSKASMT 60
DB	1	MMSCSWFMNGTGLVEELPACQDQLGLSLSLGLVGVPGVLCYNMLLVLANIHSKASMT 60
QY	61	MPDVYFVNMAVAGLVTSALAPVHLGPPSSRMLMSVGEVHVVALQIPFVSSLVAMYST 120
DB	61	MPDVYFVNMAVAGLVTSALAPVHLGPPSSRMLMSVGEVHVVALQIPFVSSLVAMYST 120
QY	121	ALLSLHYIERALPRTYMASVYTRHVCVWGAGALLTSPSSLLFTYICSHVSTRALBECAR 180
DB	121	ALLSLHYIERALPRTYMASVYTRHVCVWGAGALLTSPSSLLFTYICSHVSTRALBECAR 180
QY	181	MONAENADATLVEIGVVPALATLVALLVLSRVREDTPLDRDTGLBPSAHLVAATVC 240
DB	181	MONAENADATLVEIGVVPALATLVALLVLSRVREDTPLDRDTGLBPSAHLVAATVC 240

QY 241 TQGLWTPHYLILGHTVITISRGKPYDAHYLGHLHFVKDPSKLLAFSSSFVTPLLRYNM 300
DB 241 TQGLWTPHYLILGHTVITISRGKPYDAHYLGHLHFVKDPSKLLAFSSSFVTPLLRYNM 300
QY 301 QSFPSKIQRLMKKLPCCGRHCSPDHMGVQVLA 333
DB 301 QSFPSKIQRLMKKLPCCGRHCSPDHMGVQVLA 333

RESULT 2

US-08-480-994-38
Sequence 38, Application US/08480994
Patent No. 5834248
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,994
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-994-38

Query Match 15.6%; Score 271; DB 2; Length 375;

Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACODLQGLSLSLGLVGVGVGLCNALLVLANLSKASKMTMPDVYFV 67
DB 44 NGTG---ELSEHQOYVIGL-FLSCLYTFIFPIGVGNILIVNLSFPEKMTIDPLVFI 99
QY 68 NNAVAGLVLSALAPVHLGLPSSRWALMSVGGEVH-----VALQIPF-NVSSIVAMYSTA 121
DB 100 NNAVAGLVLSALAPVHLGLPSSRWALMSVGGEVH-----VALQIPF-NVSSIVAMYSTA 145
QY 122 L-----ISLDHYIRPALPRYTMASVNTNRH---VCGFWGAGALLTSBSSLLFYICSHVST 173
DB 146 FPLTWMSPDRIYI--ALARAARCSLIFRTKHARLSGGLIW---MASVSATLVPTAVHLQH 200

QY 174 RALECAKMONAEEADATLVFVGVV--ALATLVALLSRVREDTPIDRDYGRLEP-- 229
DB 201 TQEGCCFPADVAVQVQWLEVTLLGFIVPALIGLCYSLIVAVLRAH-----RRHG-LRPRR 254
QY 230 -SARLLVAVTCTQFGLWTPHYLILGHTVITISRGK-----VDAAHYGLHLFVKD 279
DB 255 QALREMLILAVLVFVFCWLPENVFISVH--LIQRTPGAAPCKQSFRAHAPV-TGHIIV-- 309

QY 280 PSKLLAFSSSFVTPLLRYNMNOSPSPKIQRLMKK---LPCGRHCSPDHMGVQVLA 332
DB 310 -NLAFFSNSCLNPILYSFPLGTFPRDKLRLYEQKTNLPALDRFC--HAALKAVI 360

RESULT 3

US-08-616-844-38
Sequence 38, Application US/08616844
Patent No. 5849578
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A.
TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,844
FILING DATE: 15-MAR-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,654
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-616-844-38

Query Match 15.6%; Score 271; DB 2; Length 375;

Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACODLQGLSLSLGLVGVGVGLCNALLVLANLSKASKMTMPDVYFV 67
DB 44 NGTG---ELSEHQOYVIGL-FLSCLYTFIFPIGVGNILIVNLSFPEKMTIDPLVFI 99
QY 68 NNAVAGLVLSALAPVHLGLPSSRWALMSVGGEVH-----VALQIPF-NVSSIVAMYSTA 121

Db 100 NLAVADILVADSLIEVF-----NLHERYDIAVLCTFMSLFLRYNMVSSV 145
QY 122 L-----LSLDHYIERALPRTMASVYNTRH-----VCGFWMGALLTSSSLFYICSHVST 173
Db 146 FLTWSFDRYI--ALARMRCSLFRTKHARLSCGLIW--MASVATLVPTAYHLOH 200
QY 174 RALECAGMNAEADATLVFIQVVP--ALATYVALVLSRVREDPTLDRDGRLEP-- 229
Db 201 TDEACCFADVRVQWLEVLIGFIVPFAIIGLCYSLIVRVLYRAH-----RHKG-LRPR 254
QY 230 -SAHRLVATVCTQFGIMTPTHYLLIGHTVIISRGK-----VDANYLGILHFVKD 279
Db 255 QKALRMILAVLVFVFCWLPENVFISVH--LLORTQGAAPCKQSPFHAHPL-TGHIV-- 309
QY 280 FSKLIAPSSSFVPTLYRYMNSPFSKLORLMKK---LPCGDRHCSPDHMGVOVL 332
Db 310 --NLAAFSNSCLNPLIYSFLGETFRDRLRYIEQKTNLPALDRFC--HAAKXAVI 360
RESULT 4
US-08-599-654-38
Sequence 38, Application US/08599654
Patent No. 5882925
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,654
FILING DATE: 09-FEB-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-041
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-599-654-38
Query Match 15.6%; Score 271; DB 2; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;
QY 8 NGTGLVEELPACDQLGLSLSLGLVGVGVGLCYNALLVLANHKSASMTMPDVFV 67

Db 44 NGTG---ELSEHQOYVIGL-FLSCLYTIFFLPFGVGNILIVNLSFRKMTIPDLXYI 99
QY 68 NNAVAGLVLSALAPVHLLEGPPSRMALMSVGGEVH-----VAQIIP-ANSSLVAMSTA 121
Db 100 NLAVADILVADSLIEVF-----NLHERYDIAVLCTFMSLFLRYNMVSSV 145
QY 122 L-----LSLDHYIERALPRTMASVYNTRH-----VCGFWMGALLTSSSLFYICSHVST 173
Db 146 FLTWSFDRYI--ALARMRCSLFRTKHARLSCGLIW--MASVATLVPTAYHLOH 200
QY 174 RALECAGMNAEADATLVFIQVVP--ALATYVALVLSRVREDPTLDRDGRLEP-- 229
Db 201 TDEACCFADVRVQWLEVLIGFIVPFAIIGLCYSLIVRVLYRAH-----RHKG-LRPR 254
QY 230 -SAHRLVATVCTQFGIMTPTHYLLIGHTVIISRGK-----VDANYLGILHFVKD 279
Db 255 QKALRMILAVLVFVFCWLPENVFISVH--LLORTQGAAPCKQSPFHAHPL-TGHIV-- 309
QY 280 FSKLIAPSSSFVPTLYRYMNSPFSKLORLMKK---LPCGDRHCSPDHMGVOVL 332
Db 310 --NLAAFSNSCLNPLIYSFLGETFRDRLRYIEQKTNLPALDRFC--HAAKXAVI 360

RESULT 5
US-08-485-573-38
Sequence 38, Application US/08485573
Patent No. 5968770
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,573
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-485-573-38

Query Match 15.6%; Score 271; DB 2; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-944-423A-38

Query Match 15.6%; Score 271; DB 3; Length 375;

Best Local Similarity 27.5%; Pred. No. 6,7e-17; Mismatches 124; Indels 70; Gaps 19;

Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLNLSKASMTMPDVFV 67
DB 44 NGTG---ELSEHQYVIGL-FLSCLVITPLFPIGFGVNNILLVNLSFRKMTIPDLXYI 99
QY 68 NMAVAGVLSALAPVHLGLPSSRMALMSVGSEVH-----VALQIP-NVSLVAMYSTA 121
DB 100 NLAVADLILVADSLIEVF-----NLHERYDIAVLCTFMSFLRVNMYSSV 145
QY 122 L-----LSLHYIERALPRTMASVYNTRH-----VCGFWGAGALLTSFSLFYCSHVT 173
DB 146 FLITWMSFDRIYI--ALANRACSLFRTKHARUSCGLIW--MASVATLVPTAVHLQH 200
QY 174 RALECAMQNAEADATLVFVGYVP--ALATYALVLSRVREDTPLDRDGRLEP-- 229
DB 201 TDEACCFADVREVQMLEVTLGFIVPFAITGLCYSLIVRVLAH-----RHNG-LRPR 254
QY 230 -SAHRLVATVCTQFGMLTPHYLLIGHYIISRGK-----VDHYLGILHFVKD 279
DB 255 OKALRMILAVLVFVCMLEENVFISVH--LQRTQGAAPCKOSFHAHPL-TGHIV-- 309
QY 280 FSLILAFSSFVPLIYRVNNSFPSKLOLMKK---LPCGDHRCSPDHNGVOYL 332
DB 310 --NLAFNSCLNPLIYSFLGETFRDLRYIBOKTNLPALDRFC--HAALKAVI 360

RESULT 8

US-08-925-743-38
Sequence 38, Application US/08925743

Patent No. 6034538

GENERAL INFORMATION:

APPLICANT: FALB, DEAN A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/925,743

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/485,573

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7853-032

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-925-743-38

Query Match 15.6%; Score 271; DB 3; Length 375;

Best Local Similarity 27.5%; Pred. No. 6,7e-17; Mismatches 124; Indels 70; Gaps 19;

Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLNLSKASMTMPDVFV 67
DB 44 NGTG---ELSEHQYVIGL-FLSCLVITPLFPIGFGVNNILLVNLSFRKMTIPDLXYI 99
QY 68 NMAVAGVLSALAPVHLGLPSSRMALMSVGSEVH-----VALQIP-NVSLVAMYSTA 121
DB 100 NLAVADLILVADSLIEVF-----NLHERYDIAVLCTFMSFLRVNMYSSV 145
QY 122 L-----LSLHYIERALPRTMASVYNTRH-----VCGFWGAGALLTSFSLFYCSHVT 173
DB 146 FLITWMSFDRIYI--ALANRACSLFRTKHARUSCGLIW--MASVATLVPTAVHLQH 200
QY 174 RALECAMQNAEADATLVFVGYVP--ALATYALVLSRVREDTPLDRDGRLEP-- 229
DB 201 TDEACCFADVREVQMLEVTLGFIVPFAITGLCYSLIVRVLAH-----RHNG-LRPR 254
QY 230 -SAHRLVATVCTQFGMLTPHYLLIGHYIISRGK-----VDHYLGILHFVKD 279
DB 255 OKALRMILAVLVFVCMLEENVFISVH--LQRTQGAAPCKOSFHAHPL-TGHIV-- 309
QY 280 FSLILAFSSFVPLIYRVNNSFPSKLOLMKK---LPCGDHRCSPDHNGVOYL 332
DB 310 --NLAFNSCLNPLIYSFLGETFRDLRYIBOKTNLPALDRFC--HAALKAVI 360

RESULT 9

US-08-944-496-38
Sequence 38, Application US/08944496

Patent No. 6124433

GENERAL INFORMATION:

APPLICANT: FALB, DEAN A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/944,496

FILING DATE: 06-OCT-1997

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/599,654

FILING DATE: 09-FEB-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/485,573

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/386,844

FILING DATE: 10-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: CORUZZI, LAURA A

```
/
/   REGISTRATION NUMBER: 30,742
/   REFERENCE/DOCKET NUMBER: 7853-104
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: (212) 790-9090
/   TELEFAX: (212) 869-8864
/   TELEEX: 66141 PENNIE
/   INFORMATION FOR SEQ ID NO: 38:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 375 amino acids
/   TYPE: amino acid
/   STRANDEDNESS:
/   TOPOLOGY: unknown
/   MOLECULE TYPE: protein
/   US-08-944-496-38

Query Match      15.6%; Score 271; DB 3; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACODLQGLSLSLGLGVGVPGVGLCNALLVLANLHSAKSMTPDVFV 67
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 44 NGTG--ELSEHQOYVIGL-FLSCLYTFIFPIGFGVGNILIVNISFRKMTIPDLVFI 99
/
/   68 NNAVAGLVLSALAPVHLLGPPSSRWALMSVGGEVH-----VALQIPF-NVSLVAMYSTA 121
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 100 NLAVALDLIVADSLIEVF-----NLHERRYDIAVLCTFMSLFLRVNMYSSV 145
/
/   QY 122 L-----LSLDHYTERALPRTYMASVNTRH---VGFVWGALLTSFSSILFYICSHYST 173
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 146 FFLTWMSFDRIY--ALAFAMRCSLFRTHARLSGLIW---MASVATLVPTAVHLQH 200
/
/   QY 174 RALECAKQNAEADATLVFIQVVP--ALATLYALVILSRVRDTPLDRTGRLEP-- 229
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 201 TDEACFCFADRVREOVWLEVTIGFIVPFAIIGCYSLIVRYLVRAH-----RHKG-LRRRR 254
/
/   QY 230 -SAHRLIVATVCTQFGMLTPHYLLILGHTVILSRGP-----VDHYIGLLHFVYD 279
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 255 QKALRMILAVLVFVFCWLPENVFISVH--LLQRTQPGAPCKQSFRAHPL-TGHIV-- 309
/
/   QY 280 FSKLAFSSSFVTPILLYRYNQSFPSTKQLRMK---LPCDRHSGPDMGVQOVL 332
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 310 --NLAAFNSCLNPLIYSFLGETFRDKLRLVTEQKTNLPALDRFC--HAALKAVI 360

RESULT 10
US-08-925-767-38
/   Sequence 38, Application US/08925767
/   Patent No. 6225084
/   GENERAL INFORMATION:
/   APPLICANT: FALB, DEAN A.
/   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
/   TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
/   NUMBER OF SEQUENCES: 38
/   CORRESPONDENCE ADDRESS:
/   ADDRESSEE: PENNIE & EDMONDS
/   STREET: 1155 Avenue of the Americas
/   CITY: New York
/   STATE: New York
/   COUNTRY: USA
/   ZIP: 10036-2711
/   COMPUTER READABLE FORM:
/   MEDIUM TYPE: Floppy disk
/   COMPUTER: IBM PC compatible
/   OPERATING SYSTEM: PC-DOS/MS-DOS
/   SOFTWARE: PatentIn Release #1.0, Version #1.30
/   CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/08/925,767
/   FILING DATE: 09-SEPT-1997
/   CLASSIFICATION: 514
/   PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: US 08/485,573
/   FILING DATE: 07-JUN-1995
/   PRIOR APPLICATION DATA:
```

```
/
/   APPLICATION NUMBER: US 08/386,844
/   FILING DATE: 10-FEB-1995
/   ATTORNEY/AGENT INFORMATION:
/   NAME: Coruzzi, Laura A.
/   REGISTRATION NUMBER: 30,742
/   REFERENCE/DOCKET NUMBER: 7853-097
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: (212) 790-9090
/   TELEFAX: (212) 869-8864
/   TELEEX: 66141 PENNIE
/   INFORMATION FOR SEQ ID NO: 38:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 375 amino acids
/   TYPE: amino acid
/   STRANDEDNESS:
/   TOPOLOGY: unknown
/   MOLECULE TYPE: protein
/   US-08-925-767-38
```

```
Query Match      15.6%; Score 271; DB 3; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;
```

```
QY 8 NGTGLVEELPACODLQGLSLSLGLGVGVPGVGLCNALLVLANLHSAKSMTPDVFV 67
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 44 NGTG--ELSEHQOYVIGL-FLSCLYTFIFPIGFGVGNILIVNISFRKMTIPDLVFI 99
/
/   68 NNAVAGLVLSALAPVHLLGPPSSRWALMSVGGEVH-----VALQIPF-NVSLVAMYSTA 121
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 100 NLAVALDLIVADSLIEVF-----NLHERRYDIAVLCTFMSLFLRVNMYSSV 145
/
/   QY 122 L-----LSLDHYTERALPRTYMASVNTRH---VGFVWGALLTSFSSILFYICSHYST 173
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 146 FFLTWMSFDRIY--ALAFAMRCSLFRTHARLSGLIW---MASVATLVPTAVHLQH 200
/
/   QY 174 RALECAKQNAEADATLVFIQVVP--ALATLYALVILSRVRDTPLDRTGRLEP-- 229
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 201 TDEACFCFADRVREOVWLEVTIGFIVPFAIIGCYSLIVRYLVRAH-----RHKG-LRRRR 254
/
/   QY 230 -SAHRLIVATVCTQFGMLTPHYLLILGHTVILSRGP-----VDHYIGLLHFVYD 279
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 255 QKALRMILAVLVFVFCWLPENVFISVH--LLQRTQPGAPCKQSFRAHPL-TGHIV-- 309
/
/   QY 280 FSKLAFSSSFVTPILLYRYNQSFPSTKQLRMK---LPCDRHSGPDMGVQOVL 332
/   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 310 --NLAAFNSCLNPLIYSFLGETFRDKLRLVTEQKTNLPALDRFC--HAALKAVI 360
```

```
RESULT 11
US-09-170-496D-194
/   Sequence 194, Application US/09170496D
/   Patent No. 6555339
/   GENERAL INFORMATION:
/   APPLICANT: Behan, Dominic P.
/   TITLE OF INVENTION: Chalmers, Derek T.
/   APPLICANT: Liaw, Chen W.
/   TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
/   TITLE OF INVENTION: Receptors
/   FILE REFERENCE: AREN-0040
/   CURRENT APPLICATION NUMBER: US/09/170,496D
/   NUMBER OF SEQ ID NOS: 294
/   SOFTWARE: PatentIn version 3.1
/   SEQ ID NO 194
/   LENGTH: 375
/   TYPE: PRT
/   ORGANISM: Homo sapiens
/   US-09-170-496D-194
```

```
Query Match      15.5%; Score 269; DB 4; Length 375;
Best Local Similarity 27.4%; Pred. No. 1e-16;
Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;
```


Query Match 15.4%; Score 267; DB 4; Length 375;
 Best Local Similarity 27.4%; Pred. No. 1.6e-16;

Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;

```

QY      8 NCGGLVEELPACODLQGLSLSLGLVGVPGLCYNALLVLANLHKSASGTMPPDYEV 67
DB      44 NGTG---ELSEHQOYVIGL-FLSCLYTIFFLPFGVGNILLIVNISFEKRTIIDLYFI 99
QY      68 NNAVAGLVLSALAPVHLGPPSSRWALMSVGEVVALQIPNNVSLVAMYSTAL----L 123
DB      100 NLAVADLLIVADSLLEVFNLHERRYDIIVLCTFMSLFLO-----VNMYSVFFLTMM 151
QY      124 SLDHIERALPRTYASVYNTNH---VCGFVMGALLTSFSSLLFYICSHVSTRALECA 179
DB      152 SFDRYI--ALARAMRCSLFRTHARLSGGLIM--MASVSATIVPFAVHLQHTDEACF 206
QY      180 KQONAEADATLVFIGVVP--ALATLYALVLSRVREDDTPILDRTGRLEP---SAHRL 234
DB      207 CPADVREVQMLEVTIGFIVFPFAIIGLCYSLIVRLVRAH-----RHRG-LRPRQKALRM 260
QY      235 LVATVCTQFGIMTPHYLLIGHTVIISRGR-----VDANYIGLHFPVKDFSKLLA 285
DB      261 ILAVVLVFPVCWLPENVFISVH--LLQRTQGAAPCKOSFRHAHPL-TGHIV---NLAA 313
QY      286 FSSSFVTPLLYRYNMOSFPSKLORLMKK---LPCGDRHCSPDHMGVOOVL 332
DB      314 FSNSCINPLIYSFLGETFRDKRLTYIEOKTNLPALNRFPC---HAALKAVI 360

```

RESULT 15

US-09-170-496D-56

/ Sequence 56, Application US/09170496D

/ Patent No. 6555339

/ GENERAL INFORMATION:

/ APPLICANT: Behan, Dominic P.

/ APPLICANT: Chalmers, Derek T.

/ APPLICANT: Liaw, Chen W.

/ TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-

/ FILE REFERENCE: AREN-0040

/ CURRENT APPLICATION NUMBER: US/09/170,496D

/ NUMBER OF SEQ ID NOS: 294

/ SOFTWARE: PatentIn version 3.1

/ SEQ ID NO 56

/ LENGTH: 375

/ TYPE: PRT

/ ORGANISM: Homo sapiens

US-09-170-496D-56

Query Match 15.4%; Score 267; DB 4; Length 375;
 Best Local Similarity 27.4%; Pred. No. 1.6e-16;

Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;

```

QY      8 NCGGLVEELPACODLQGLSLSLGLVGVPGLCYNALLVLANLHKSASGTMPPDYEV 67
DB      44 NGTG---ELSEHQOYVIGL-FLSCLYTIFFLPFGVGNILLIVNISFEKRTIIDLYFI 99
QY      68 NNAVAGLVLSALAPVHLGPPSSRWALMSVGEVVALQIPNNVSLVAMYSTAL----L 123
DB      100 NLAVADLLIVADSLLEVFNLHERRYDIIVLCTFMSLFLO-----VNMYSVFFLTMM 151
QY      124 SLDHIERALPRTYASVYNTNH---VCGFVMGALLTSFSSLLFYICSHVSTRALECA 179
DB      152 SFDRYI--ALARAMRCSLFRTHARLSGGLIM--MASVSATIVPFAVHLQHTDEACF 206
QY      180 KQONAEADATLVFIGVVP--ALATLYALVLSRVREDDTPILDRTGRLEP---SAHRL 234
DB      207 CPADVREVQMLEVTIGFIVFPFAIIGLCYSLIVRLVRAH-----RHRG-LRPRQKALRM 260
QY      235 LVATVCTQFGIMTPHYLLIGHTVIISRGR-----VDANYIGLHFPVKDFSKLLA 285
DB      261 ILAVVLVFPVCWLPENVFISVH--LLQRTQGAAPCKOSFRHAHPL-TGHIV---NLAA 313

```

```

QY      286 FSSSFVTPLLYRYNMOSFPSKLORLMKK---LPCGDRHCSPDHMGVOOVL 332
DB      314 FSNSCINPLIYSFLGETFRDKRLTYIEOKTNLPALNRFPC---HAALKAVI 360

```

Search completed: December 14, 2004, 20:28:12
 Job time : 16.5335 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:23:14 ; Search time 46.6006 Seconds
(Without alignments)
2552.337 Million cell updates/sec

Title: US-09-995-225B-2

Perfect score: 1733
Sequence: 1 MMSCSMFNMGVLVEELPACQ.....LPQGDHRCSPDHMGVQVLA 333

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*

18: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep:*

19: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep:*

20: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the distribution,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1733	100.0	333	9	US-09-995-225-2
2	1733	100.0	333	10	US-09-791-932-114
3	1733	100.0	333	10	US-09-990-940-4
4	1733	100.0	333	10	US-09-995-225-2
5	1733	100.0	333	14	US-10-165-844-13
6	1733	100.0	333	14	US-10-101-148-3
7	1733	100.0	333	14	US-10-292-798-1016
8	1733	100.0	370	15	US-10-296-115-1174
9	1733	100.0	366	14	US-10-017-161-11202
10	1726	99.6	333	9	US-09-895-686-3
11	1726	99.6	333	10	US-09-791-932-91
12	1717	99.1	333	14	US-10-225-567A-660
13	271	15.6	375	9	US-09-371-900-38

14	271	15.6	375	9	US-09-924-417-59	Sequence 59, Appl
15	271	15.6	375	9	US-09-970-820-38	Sequence 38, Appl
16	271	15.6	375	9	US-09-986-718-38	Sequence 38, Appl
17	271	15.6	375	14	US-10-186-950-38	Sequence 38, Appl
18	271	15.6	375	15	US-10-653-872-59	Sequence 59, Appl
19	269	15.5	375	14	US-10-251-385-194	Sequence 194, Appl
20	269	15.5	375	14	US-10-251-385-202	Sequence 202, Appl
21	267	15.4	375	14	US-10-251-385-56	Sequence 56, Appl
22	267	15.4	375	14	US-10-251-385-70	Sequence 70, Appl
23	267	15.4	375	14	US-10-225-567A-92	Sequence 92, Appl
24	210.5	12.1	355	14	US-10-237-563-37	Sequence 37, Appl
25	208.5	12.0	355	14	US-10-237-563-34	Sequence 34, Appl
26	207.5	12.0	355	14	US-10-237-563-36	Sequence 36, Appl
27	201.5	11.6	355	14	US-10-237-563-35	Sequence 35, Appl
28	201	11.6	355	14	US-10-237-563-25	Sequence 29, Appl
29	201	11.6	355	14	US-10-237-563-30	Sequence 30, Appl
30	200	11.5	354	14	US-10-411-284-11	Sequence 11, Appl
31	200	11.5	355	10	US-09-893-512-10	Sequence 10, Appl
32	200	11.5	355	14	US-10-237-563-27	Sequence 27, Appl
33	200	11.5	355	14	US-10-237-563-28	Sequence 28, Appl
34	200	11.5	355	14	US-10-237-563-31	Sequence 31, Appl
35	200	11.5	355	17	US-10-791-592-8	Sequence 8, Appl
36	200	11.5	355	17	US-10-791-166-8	Sequence 8, Appl
37	200	11.5	360	13	US-10-087-192-882	Sequence 882, Appl
38	200	11.5	360	14	US-10-099-007A-16	Sequence 16, Appl
39	200	11.5	360	14	US-10-225-567A-386	Sequence 386, Appl
40	200	11.5	360	14	US-10-239-423-77	Sequence 77, Appl
41	200	11.5	360	14	US-10-101-148-4	Sequence 4, Appl
42	200	11.5	399	15	US-10-334-143-7	Sequence 7, Appl
43	200	11.5	399	15	US-10-276-774-2394	Sequence 2394, Ap
44	193.5	11.2	350	9	US-09-782-980-83	Sequence 83, Appl
45	193.5	11.2	350	9	US-09-884-430-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-995-225-2
Sequence 2, Appl
Publication No. US20020193584A1
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Dang, Hong T.
APPLICANT: Lowitz, Kevin P.
TITLE OF INVENTION: Pride, Cameron
TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human
FILE REFERENCE: ARN-0308
CURRENT APPLICATION NUMBER: US/09/995,225
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917

;; PRIOR FILING DATE: 2001-05-14
;; PRIOR APPLICATION NUMBER: 60/309,208
;; PRIOR FILING DATE: 2001-07-31
;; NUMBER OF SEQ ID NOS: 67
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 333
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: No. US20020193584A1el Sequence
US-09-995-225-2

Query Match 100.0%; Score 1733; DB 9; Length 333;
Best Local Similarity 100.0%; Pred. No. 9, 8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSGWFNGTGLVEELPACODIQLGSLSLGLVVGVPVGLCYNALLVLANLHSAKAMT 60
DB 1 MMSGWFNGTGLVEELPACODIQLGSLSLGLVVGVPVGLCYNALLVLANLHSAKAMT 60
QY 61 MEDVYFVNMAVAGLVLSALAPVHLLGPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MEDVYFVNMAVAGLVLSALAPVHLLGPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSDHYTERALPRTYMASVYNTRHVCGFWGALLTSFSSILFYICSHVSTRALBCKA 180
DB 121 ALLSDHYTERALPRTYMASVYNTRHVCGFWGALLTSFSSILFYICSHVSTRALBCKA 180
QY 181 MONAADAATLVFVIGYVPALATLVALLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
DB 181 MONAADAATLVFVIGYVPALATLVALLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
QY 241 TQFGMTPHYLLILGHTVYISRGKPYDAHYGLHFVDFSKLAFSSSFVTPLLYRYMN 300
DB 241 TQFGMTPHYLLILGHTVYISRGKPYDAHYGLHFVDFSKLAFSSSFVTPLLYRYMN 300
QY 301 QSFPSKIQRLMKCLPCGDRHCSPDHMGVQOYLA 333
DB 301 QSFPSKIQRLMKCLPCGDRHCSPDHMGVQOYLA 333

RESULT 2
US-09-791-932-114
;; Sequence 114, Application US/09791932
;; Publication No. US20030003451A1
;; GENERAL INFORMATION:
;; APPLICANT: Vogel, Gabriel
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Hiebsch, Ronald R.
;; APPLICANT: Lind, Peter
;; APPLICANT: Kayes, Paul S.
;; APPLICANT: Ruff, Valerie
;; APPLICANT: Hufe, Rita M.
;; APPLICANT: Wood, Linda S.
;; TITLE OF INVENTION: No. US20030003451A1el G Protein-Coupled Receptors Cross-Referen
;; FILE REFERENCE: 00325, US1
;; CURRENT APPLICATION NUMBER: US/09/791, 932
;; CURRENT FILING DATE: 2001-02-23
;; PRIOR APPLICATION NUMBER: 60/184,305
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/184,304
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/184,303
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/184,397
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/184,247
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/188,880
;; PRIOR FILING DATE: 2000-03-13
;; PRIOR APPLICATION NUMBER: 60/217,369
;; PRIOR FILING DATE: 2000-07-11

;; PRIOR APPLICATION NUMBER: 60/217,370
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/218,492
;; PRIOR FILING DATE: 2000-07-20
;; PRIOR APPLICATION NUMBER: 60/186,810
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/188,064
;; PRIOR FILING DATE: 2000-03-09
;; PRIOR APPLICATION NUMBER: 60/186,457
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: 60/213,861
;; PRIOR FILING DATE: 2000-06-23
;; PRIOR APPLICATION NUMBER: 60/194,344
;; PRIOR FILING DATE: 2000-04-03
;; PRIOR APPLICATION NUMBER: 60/218,337
;; PRIOR FILING DATE: 2000-07-14
;; NUMBER OF SEQ ID NOS: 184
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 114
;; LENGTH: 333
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-791-932-114

Query Match 100.0%; Score 1733; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 9, 8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSGWFNGTGLVEELPACODIQLGSLSLGLVVGVPVGLCYNALLVLANLHSAKAMT 60
DB 1 MMSGWFNGTGLVEELPACODIQLGSLSLGLVVGVPVGLCYNALLVLANLHSAKAMT 60
QY 61 MEDVYFVNMAVAGLVLSALAPVHLLGPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MEDVYFVNMAVAGLVLSALAPVHLLGPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSDHYTERALPRTYMASVYNTRHVCGFWGALLTSFSSILFYICSHVSTRALBCKA 180
DB 121 ALLSDHYTERALPRTYMASVYNTRHVCGFWGALLTSFSSILFYICSHVSTRALBCKA 180
QY 181 MONAADAATLVFVIGYVPALATLVALLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
DB 181 MONAADAATLVFVIGYVPALATLVALLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
QY 241 TQFGMTPHYLLILGHTVYISRGKPYDAHYGLHFVDFSKLAFSSSFVTPLLYRYMN 300
DB 241 TQFGMTPHYLLILGHTVYISRGKPYDAHYGLHFVDFSKLAFSSSFVTPLLYRYMN 300
QY 301 QSFPSKIQRLMKCLPCGDRHCSPDHMGVQOYLA 333
DB 301 QSFPSKIQRLMKCLPCGDRHCSPDHMGVQOYLA 333

RESULT 3
US-09-990-940-4
;; Sequence 4, Application US/09990940
;; Publication No. US20030027252A1
;; GENERAL INFORMATION:
;; APPLICANT: Tian, Hui
;; APPLICANT: Zhao, Jiayang
;; APPLICANT: Chen, Jin-Long
;; APPLICANT: Cutler, Gene
;; APPLICANT: An, Songzhu
;; APPLICANT: Dai, Kang
;; APPLICANT: Gupta, Jamila S.
;; APPLICANT: Tularik Inc.
;; TITLE OF INVENTION: No. US20030027252A1el Receptors
;; FILE REFERENCE: 018781-007410US
;; CURRENT APPLICATION NUMBER: US/09/990, 940
;; CURRENT FILING DATE: 2001-11-21
;; PRIOR APPLICATION NUMBER: US 60/252,841
;; PRIOR FILING DATE: 2000-11-22
;; PRIOR APPLICATION NUMBER: US 60/257,636


```
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 60/261,377
/ PRIOR FILING DATE: 2001-01-12
/ PRIOR APPLICATION NUMBER: US 60/279,554
/ PRIOR FILING DATE: 2001-03-28
/ PRIOR APPLICATION NUMBER: US 60/280,696
/ PRIOR FILING DATE: 2001-03-29
/ NUMBER OF SEQ ID NOS: 54
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 333
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR339
US-09-990-940-4

Query Match          100.0%; Score 1733; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 9,8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWFGTGLVEBELPACODDQLGLSLSLGLVGVPGVGLCYNALLVLANHKSASMT 60
DB 1 MMSCSWFGTGLVEBELPACODDQLGLSLSLGLVGVPGVGLCYNALLVLANHKSASMT 60

QY 61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGSVHVALQIPFVSSLVAMYST 120
DB 61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGSVHVALQIPFVSSLVAMYST 120

QY 121 ALLSDHYIERALPRTYMASVYNTRVHVCQFVWGALLTSFSSLLFYICSHVSTRALBEC 180
DB 121 ALLSDHYIERALPRTYMASVYNTRVHVCQFVWGALLTSFSSLLFYICSHVSTRALBEC 180

QY 181 MONEAADATLVFVIGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
DB 181 MONEAADATLVFVIGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240

QY 241 TORGLWTPHYLLLGHTVVIISRGKPVDAHYLGHLFPKDSKLLAFSSFTVTPLYRYNM 300
DB 241 TORGLWTPHYLLLGHTVVIISRGKPVDAHYLGHLFPKDSKLLAFSSFTVTPLYRYNM 300

QY 301 QSFPSKLGRLMKKLPCGDRHCSPDHMGVQVLA 333
DB 301 QSFPSKLGRLMKKLPCGDRHCSPDHMGVQVLA 333

RESULT 4
US-09-995-225-2
/ Sequence 2, Application US/09995225
/ Publication No. US20030139588A9
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Ruoping
/ APPLICANT: Chu, Zhi Liang
/ APPLICANT: Dang, Huong T.
/ APPLICANT: Lowitz, Kevin P.
/ APPLICANT: Prider, Cameron
/ TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
/ FILE REFERENCE: AREN-0308
/ CURRENT APPLICATION NUMBER: US/09/995,225
/ PRIOR FILING DATE: 2001-11-26
/ PRIOR APPLICATION NUMBER: 09/170,496
/ PRIOR FILING DATE: 1998-10-13
/ PRIOR APPLICATION NUMBER: PCT/US99/23938
/ PRIOR FILING DATE: 1998-10-13
/ PRIOR APPLICATION NUMBER: 60/253,404
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/255,366
/ PRIOR FILING DATE: 2000-12-12
/ PRIOR APPLICATION NUMBER: 60/270,286
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/282,365
/ PRIOR FILING DATE: 2001-04-06
```

```
/ PRIOR APPLICATION NUMBER: 60/270,266
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/282,032
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,358
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,356
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/290,917
/ PRIOR FILING DATE: 2001-05-14
/ PRIOR APPLICATION NUMBER: 60/309,208
/ PRIOR FILING DATE: 2001-07-31
/ NUMBER OF SEQ ID NOS: 67
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 333
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: No. US20030139588A9e1 Sequence
US-09-995-225-2

Query Match          100.0%; Score 1733; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 9,8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWFGTGLVEBELPACODDQLGLSLSLGLVGVPGVGLCYNALLVLANHKSASMT 60
DB 1 MMSCSWFGTGLVEBELPACODDQLGLSLSLGLVGVPGVGLCYNALLVLANHKSASMT 60

QY 61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGSVHVALQIPFVSSLVAMYST 120
DB 61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGSVHVALQIPFVSSLVAMYST 120

QY 121 ALLSDHYIERALPRTYMASVYNTRVHVCQFVWGALLTSFSSLLFYICSHVSTRALBEC 180
DB 121 ALLSDHYIERALPRTYMASVYNTRVHVCQFVWGALLTSFSSLLFYICSHVSTRALBEC 180

QY 181 MONEAADATLVFVIGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
DB 181 MONEAADATLVFVIGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240

QY 241 TORGLWTPHYLLLGHTVVIISRGKPVDAHYLGHLFPKDSKLLAFSSFTVTPLYRYNM 300
DB 241 TORGLWTPHYLLLGHTVVIISRGKPVDAHYLGHLFPKDSKLLAFSSFTVTPLYRYNM 300

QY 301 QSFPSKLGRLMKKLPCGDRHCSPDHMGVQVLA 333
DB 301 QSFPSKLGRLMKKLPCGDRHCSPDHMGVQVLA 333

RESULT 5
US-10-165-844-13
/ Sequence 13, Application US/10165844
/ Publication No. US20030017539A1
/ GENERAL INFORMATION:
/ APPLICANT: Gluckemann, Maria Alexandra
/ APPLICANT: Hodge, Martin R.
/ APPLICANT: Hunter, John J.
/ APPLICANT: Rudolph-Owen, Laura
/ APPLICANT: Welch, Nadine S.
/ APPLICANT: Siles-Santiego, Immaculada
/ TITLE OF INVENTION: Novel Nucleic Acid Sequences Encoding
/ FILE REFERENCE: 35800/248302
/ CURRENT APPLICATION NUMBER: US/10/165,844
/ PRIOR FILING DATE: 2002-06-07
/ PRIOR APPLICATION NUMBER: US 09/088,857
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: US 09/324,465
/ PRIOR FILING DATE: 1999-06-02
/ PRIOR APPLICATION NUMBER: US 09/464,685
/ PRIOR FILING DATE: 1999-12-16
```

PRIOR APPLICATION NUMBER: US 09/741,783
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: US 09/145,745
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: US 09/383,745
PRIOR FILING DATE: 1999-08-26
PRIOR APPLICATION NUMBER: US 09/234,923
PRIOR FILING DATE: 1999-01-21
PRIOR APPLICATION NUMBER: US 09/340,880
PRIOR FILING DATE: 1999-06-28
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-10-165-844-13

Query Match 100.0%; Score 1733; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 9, 8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWFGNGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNALLVLANIHSKASMT 60
DB 1 MMSCSWFGNGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNALLVLANIHSKASMT 60
QY 61 MPDYVFVNMAVAGVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
DB 61 MPDYVFVNMAVAGVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
QY 121 ALLSLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTFSLSLFYICSHVSTRALECAK 180
DB 121 ALLSLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTFSLSLFYICSHVSTRALECAK 180
QY 181 MONAERADATLVFVGVVPALATLYALVLSRVRREDTPLDRDTGRLEPSAHRLLVATVC 240
DB 181 MONAERADATLVFVGVVPALATLYALVLSRVRREDTPLDRDTGRLEPSAHRLLVATVC 240
QY 241 TOFGMTPHYLLILGHTVYIISRGKPVDAHYGLHFVDFSKLAFSSFTPLLYRYMN 300
DB 241 TOFGMTPHYLLILGHTVYIISRGKPVDAHYGLHFVDFSKLAFSSFTPLLYRYMN 300
QY 301 OSFPSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 OSFPSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 6
US-10-101-148-3
Sequence 3, Application US/10101148
Publication No. US20030187198A1
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Yeung, George
TITLE OF INVENTION: A No. US20030187198A1 Chemokine Receptor Obtained From a CDNA I
FILE REFERENCE: 28110/38317
CURRENT APPLICATION NUMBER: US/10/101,148
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin version 3.1
SEQ ID NO 3
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-10-101-148-3

Query Match 100.0%; Score 1733; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 9, 8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWFGNGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNALLVLANIHSKASMT 60
DB 1 MMSCSWFGNGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNALLVLANIHSKASMT 60

DB 1 MMSCSWFGNGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNALLVLANIHSKASMT 60
QY 61 MPDYVFVNMAVAGVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
DB 61 MPDYVFVNMAVAGVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
QY 121 ALLSLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTFSLSLFYICSHVSTRALECAK 180
DB 121 ALLSLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTFSLSLFYICSHVSTRALECAK 180
QY 181 MONAERADATLVFVGVVPALATLYALVLSRVRREDTPLDRDTGRLEPSAHRLLVATVC 240
DB 181 MONAERADATLVFVGVVPALATLYALVLSRVRREDTPLDRDTGRLEPSAHRLLVATVC 240
QY 241 TOFGMTPHYLLILGHTVYIISRGKPVDAHYGLHFVDFSKLAFSSFTPLLYRYMN 300
DB 241 TOFGMTPHYLLILGHTVYIISRGKPVDAHYGLHFVDFSKLAFSSFTPLLYRYMN 300
QY 301 OSFPSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 OSFPSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 7
US-10-292-798-1016
Sequence 1016, Application US/10292798
Publication No. US20030235833A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: GINANSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1016
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-10-292-798-1016

Query Match 100.0%; Score 1733; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 9, 8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWFGNGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNALLVLANIHSKASMT 60
DB 1 MMSCSWFGNGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNALLVLANIHSKASMT 60
QY 61 MPDYVFVNMAVAGVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
DB 61 MPDYVFVNMAVAGVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
QY 121 ALLSLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTFSLSLFYICSHVSTRALECAK 180
DB 121 ALLSLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTFSLSLFYICSHVSTRALECAK 180
QY 181 MONAERADATLVFVGVVPALATLYALVLSRVRREDTPLDRDTGRLEPSAHRLLVATVC 240
DB 181 MONAERADATLVFVGVVPALATLYALVLSRVRREDTPLDRDTGRLEPSAHRLLVATVC 240
QY 241 TOFGMTPHYLLILGHTVYIISRGKPVDAHYGLHFVDFSKLAFSSFTPLLYRYMN 300
DB 241 TOFGMTPHYLLILGHTVYIISRGKPVDAHYGLHFVDFSKLAFSSFTPLLYRYMN 300
QY 301 OSFPSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 OSFPSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333

Db 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 8
US-10-296-115-1174
Sequence 1174, Application US/10296115
Publication No. US20040053248A1
GENERAL INFORMATION:
APPLICANT: Hyseq Inc
TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
FILE REFERENCE: 784PCT
CURRENT APPLICATION NUMBER: US/10/296,115
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: US09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 1478
SEQ ID NO 1174
LENGTH: 370
TYPE: PRT
ORGANISM: Homo sapiens
US-10-296-115-1174

Query Match 100.0%; Score 1733; DB 15; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.1e-152;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWFGNGTGLVEELPACQDQLGSLSLGLVVGVPVGLCYNALLVTANLHSTRASMT 60
DB 38 MMSCSWFGNGTGLVEELPACQDQLGSLSLGLVVGVPVGLCYNALLVTANLHSTRASMT 97
QY 61 MPDYVFNMAVAGLVLSALAPVHLGPPSSRNALMSVGEVHALQIPFNVSLLVAMYST 120
DB 98 MPDYVFNMAVAGLVLSALAPVHLGPPSSRNALMSVGEVHALQIPFNVSLLVAMYST 157
QY 121 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 158 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 217
QY 181 MQNAEADATLVFVIGVVPALATLVALLSRVREDTPLDRDTGRLSPAHRLVAATVC 240
DB 218 MQNAEADATLVFVIGVVPALATLVALLSRVREDTPLDRDTGRLSPAHRLVAATVC 277
QY 241 TQFGLMTPHYLLILGHTVLIISRGKPYDAHYLGHLHFVKDFSKLLAPSSFTPLVRYNM 300
DB 278 TQFGLMTPHYLLILGHTVLIISRGKPYDAHYLGHLHFVKDFSKLLAPSSFTPLVRYNM 337
QY 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 338 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 370

RESULT 9
US-10-017-161-1202
Sequence 1202, Application US/10017161
Publication No. US20030143668A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1202
LENGTH: 386

TYPE: PRT
ORGANISM: Homo sapiens
US-10-017-161-1202

Query Match 100.0%; Score 1733; DB 14; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.2e-152;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWFGNGTGLVEELPACQDQLGSLSLGLVVGVPVGLCYNALLVTANLHSTRASMT 60
DB 54 MMSCSWFGNGTGLVEELPACQDQLGSLSLGLVVGVPVGLCYNALLVTANLHSTRASMT 113
QY 61 MPDYVFNMAVAGLVLSALAPVHLGPPSSRNALMSVGEVHALQIPFNVSLLVAMYST 120
DB 114 MPDYVFNMAVAGLVLSALAPVHLGPPSSRNALMSVGEVHALQIPFNVSLLVAMYST 173
QY 121 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 174 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 233
QY 181 MQNAEADATLVFVIGVVPALATLVALLSRVREDTPLDRDTGRLSPAHRLVAATVC 240
DB 234 MQNAEADATLVFVIGVVPALATLVALLSRVREDTPLDRDTGRLSPAHRLVAATVC 293
QY 241 TQFGLMTPHYLLILGHTVLIISRGKPYDAHYLGHLHFVKDFSKLLAPSSFTPLVRYNM 300
DB 294 TQFGLMTPHYLLILGHTVLIISRGKPYDAHYLGHLHFVKDFSKLLAPSSFTPLVRYNM 353
QY 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 354 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 386

RESULT 10
US-09-895-686-3
Sequence 3, Application US/09895686
Patent No. US2002010655A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Tang, Y. Tom
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: HUMAN GPCR PROTEINS
FILE REFERENCE: PC-0044 CIP
CURRENT APPLICATION NUMBER: US/09/895,686
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 74
SOFTWARE: PERL Program
SEQ ID NO 3
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US2002010655A1 2214673CD1
US-09-895-686-3

Query Match 99.6%; Score 1726; DB 9; Length 333;
Best Local Similarity 99.7%; Pred. No. 4.4e-152;
Matches 332; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMSCSWFGNGTGLVEELPACQDQLGSLSLGLVVGVPVGLCYNALLVTANLHSTRASMT 60
DB 1 MMSCSWFGNGTGLVEELPACQDQLGSLSLGLVVGVPVGLCYNALLVTANLHSTRASMT 113
QY 61 MPDYVFNMAVAGLVLSALAPVHLGPPSSRNALMSVGEVHALQIPFNVSLLVAMYST 120
DB 61 MPDYVFNMAVAGLVLSALAPVHLGPPSSRNALMSVGEVHALQIPFNVSLLVAMYST 173
QY 121 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 121 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180

QY 181 MONEAADATLVFIGVVPALATLVLTLSRVREDDPDLRDTGRLEPSAHLIVATVC 240
DB 181 MONEAADATLVFIGVVPALATLVLTLSRVREDDPDLRDTGRLEPSAHLIVATVC 240
QY 241 TOFGIMTPHYLLILGHVTVISRGKPVDAHVYGLLHFVDFSKLAFSSFTPLLYRYNN 300
DB 241 TOFGIMTPHYLLILGHVTVISRGKPVDAHVYGLLHFVDFSKLAFSSFTPLLYRYNN 300
QY 301 QSPFSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSPFSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 11

US-09-791-932-91
Sequence 91, Application US/09791932
Publication No. US20030003451A1
GENERAL INFORMATION:
APPLICANT: Vogeli, Gabriel
APPLICANT: Parodi, Luis A.
APPLICANT: Hiebsch, Ronald R.
APPLICANT: Lind, Peter
APPLICANT: Kayles, Paul S.
APPLICANT: Ruff, Valerie
APPLICANT: Huff, Rita M.
APPLICANT: Wood, Linda S.
TITLE OF INVENTION: No. US20030003451A1 G Protein-Coupled Receptors Cross-Referen
FILE REFERENCE: 00325 US1
CURRENT APPLICATION NUMBER: US/09/791,932
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/184,305
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,304
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,303
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,397
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,247
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/188,880
PRIOR FILING DATE: 2000-03-13
PRIOR APPLICATION NUMBER: 60/217,369
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/217,370
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/218,492
PRIOR FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: 60/186,810
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/188,064
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 60/186,457
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: 60/213,861
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/194,344
PRIOR FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: 60/218,337
PRIOR FILING DATE: 2000-07-14
NUMBER OF SEQ ID NOS: 184
SOFTWARE: Patentin version 3.0
SEQ ID NO 91
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-932-91

Query Match 99.6%; Score 1726; DB 10; Length 333;
Best Local Similarity 99.7%; Pred. No. 4,4e-152;
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MMSCSWFGTGLVEELPACODLQLGSLSLILGLVGVGVGLCYNNALVLANLHRSKASMT 60

DB 1 MMSCSWFGTGLVEELPACODLQLGSLSLILGLVGVGVGLCYNNALVLANLHRSKASMT 60
QY 61 MPDYVFVNNAAVAGLVLSALAPVHLGPPSSRWALMSVGGEEVVALQIPNVSSLVAMYST 120
DB 61 MPDYVFVNNAAVAGLVLSALAPVHLGPPSSRWALMSVGGEEVVALQIPNVSSLVAMYST 120
QY 121 ALLSLDHYTERALPRTYMASVYNTRHVCGFWGAGALLTSFSSILFYICSHVSTRALBCKA 180
DB 121 ALLSLDHYTERALPRTYMASVYNTRHVCGFWGAGALLTSFSSILFYICSHVSTRALBCKA 180
QY 181 MONEAADATLVFIGVVPALATLVLTLSRVREDDPDLRDTGRLEPSAHLIVATVC 240
DB 181 MONEAADATLVFIGVVPALATLVLTLSRVREDDPDLRDTGRLEPSAHLIVATVC 240
QY 241 TOFGIMTPHYLLILGHVTVISRGKPVDAHVYGLLHFVDFSKLAFSSFTPLLYRYNN 300
DB 241 TOFGIMTPHYLLILGHVTVISRGKPVDAHVYGLLHFVDFSKLAFSSFTPLLYRYNN 300
QY 301 QSPFSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSPFSKLORLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 12

US-10-225-567A-680
Sequence 680, Application US/10225567A
Publication No. US20030113798A1
GENERAL INFORMATION:
APPLICANT: Lifespan Biosciences
APPLICANT: Brown, Joseph P.
APPLICANT: Burnet, Glema C.
APPLICANT: Roush, Christine L.
TITLE OF INVENTION: ANTIDENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 1920-4-4
CURRENT APPLICATION NUMBER: US/10/225,567A
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/257,144
PRIOR FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 2292
SOFTWARE: Patentin version 3.1
SEQ ID NO 680
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (11)-(11)
OTHER INFORMATION: Unknown Amino Acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (117)-(117)
OTHER INFORMATION: Unknown Amino Acid
US-10-225-567A-680

Query Match 99.1%; Score 1717; DB 14; Length 333;
Best Local Similarity 99.4%; Pred. No. 3e-151;
Matches 331; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MMSCSWFGTGLVEELPACODLQLGSLSLILGLVGVGVGLCYNNALVLANLHRSKASMT 60
DB 1 MMSCSWFGTGLVEELPACODLQLGSLSLILGLVGVGVGLCYNNALVLANLHRSKASMT 60
QY 61 MPDYVFVNNAAVAGLVLSALAPVHLGPPSSRWALMSVGGEEVVALQIPNVSSLVAMYST 120
DB 61 MPDYVFVNNAAVAGLVLSALAPVHLGPPSSRWALMSVGGEEVVALQIPNVSSLVAMYST 120
QY 121 ALLSLDHYTERALPRTYMASVYNTRHVCGFWGAGALLTSFSSILFYICSHVSTRALBCKA 180
DB 121 ALLSLDHYTERALPRTYMASVYNTRHVCGFWGAGALLTSFSSILFYICSHVSTRALBCKA 180
QY 181 MONEAADATLVFIGVVPALATLVLTLSRVREDDPDLRDTGRLEPSAHLIVATVC 240
DB 181 MONEAADATLVFIGVVPALATLVLTLSRVREDDPDLRDTGRLEPSAHLIVATVC 240

Db 181 MONEAADATLVPIGVVPALATLVTLRSVRREDTDLDRGTGRLESPARLLVAVTC 240
Qy 241 TQGLMTPHYLLILGHVVIISRGKPVDAHYIGLHFPYKDPKLAFFSSFTPLLYRYM 300
Db 241 TQGLMTPHYLLILGHVVIISRGKPVDAHYIGLHFPYKDPKLAFFSSFTPLLYRYM 300
Qy 301 QSPSKLQRLMKKLPCCDRHCSPDHMGVOVLA 333
Db 301 QSPSKLQRLMKKLPCCDRHCSPDHMGVOVLA 333

RESULT 13
US-09-371-900-38
Sequence 38: Application US/09371900
Patent No. US20020137700A1
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
City: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/371.900
FILING DATE: 11-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,654
FILING DATE: 09-FEB-1996
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-371-900-38

Query Match 15.6%; Score 271; DB 9; Length 375;
Best Local Similarity 27.5%; Pred. No. 1,4e-16;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

Qy 8 NGTGLVEELPACDPLQGLSLSLGLVGVVGLCYNALVLNLHASKAMTPDYPV 67
Db 44 NGTG---ELSEHQVYIGL-PLSLTYIFLPPIGFVGNILIVNISFRERKTPDIXPI 99
Qy 68 NMAVAGIVLALAPVHLGLPPSSRWALMSVGEVH-----VALQIPF-NVSSIVAMYSTA 121
Db 100 NLAVALDILVADSLIEVF-----NLHERYDIADVCTMSIFLRVNMYSV 145

Qy 122 L---LSLDHYERIALPRTYMASVNTNRH-----VCGFVMGALLTSPSSILFYICSHVST 173
Db 146 FFLTMSFDXYI--ALAMARCSLFRTHKHAISCGLIW---MASVATLVPTAVHLQH 200
Qy 174 RALECANMONEAADATLVPIGVVP--ALATLVTLRSVRREDTDLDRGTGRLEP-- 229
Db 201 TDEACFCFADVRREVQMLLEVTLGFTVFPALIGLCSYLIVLVBAH-----RHKG-LRPR 254
Qy 230 -SAHRLVAVTCQFGMLTPHYLLILGHVVIISRGKPVDAHYIGLHFPYKDPKLAFFSSFTPLLYRYM 300
Db 255 QKALRMILAVLVFVFCWLPENVFISVH--LQRTQFGAAPCKOSFRAHPL-TGHTV-- 309
Qy 280 PSKLAFFSSFTPLLYRYMNSPPSKLQRLMKK--LPCDRHCSPDHMGVOVLA 332
Db 310 -NLAFFNSCLNPLISFLGETFRDKRLYIEKTNLPADNFC--HAALAAVI 360

RESULT 14
US-09-924-417-59
Sequence 59: Application US/09924417
Patent No. US2002014241A1
GENERAL INFORMATION:
APPLICANT: FALB, DEAN
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR
DISEASE
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
City: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patco Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/924.417
FILING DATE: 07-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/034,286
FILING DATE: 04-MAR-1998
APPLICATION NUMBER: 08/870,434
FILING DATE: 06-JUN-1997
APPLICATION NUMBER: 08/799,910
FILING DATE: 13-FEB-1997
APPLICATION NUMBER: 60/011,787
FILING DATE: 16-FEB-1996
APPLICATION NUMBER: 08/599,654
FILING DATE: 09-FEB-1996
APPLICATION NUMBER: 08/485,573
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-114-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 7909090
TELEFAX: (212) 8699741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-924-417-59

Query Match 15.6%; Score 271; DB 9; Length 375;
Best Local Similarity 27.5%; Pred. No. 1.4e-16;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACQDLDGLSLSLGLVGVPGVCYNALVLANLHSAKSTMPDYFV 67
DB 44 NGTG--ELSEHQOYVIGL-FLSCLYTIFLPFIGVGNILILVNNISFREKKTIDLYFI 99
QY 68 NNAVAGLVLSALAPVHLLGPSSRWALMSVGGEVH--VALQIPF-NVSSIVAMYSTA 121
DB 100 NIAVADLLIVADSLIEVF-----NLHERYYDIAVLCTFMSLFLRVNMYSV 145
QY 122 L-----LSLDHYTERALPRTYMASVYNTRH---VCGFWGALITSFSSLLFYICSHYST 173
DB 146 FLITWMSFDRYI--ALARAMCSLFRTHHARLSGLIM--MASVSATLVPTAVHLOH 200
QY 174 RALECAKQNAEADATLVFIGYVP--ALATVVALVLSRVREDTFLDRDTGRLEP-- 229
DB 201 TDEACFCPADVREVOVMEVTLGFIIVFPALIGLCYSLIVRVLRH-----RHKG-LRPBR 254
QY 230 -SAHRLVAVTCTOGMTTPHYLLIGHTVLIISRGK-----VDANYLGHLHFVKD 279
DB 255 OKALRMIIAVALVVFVFCWLPENVFISVH--LLQRTQGAAPCKOSFRHAPL-TGHIV-- 309
QY 260 FSKLIAFSSPYTPPLVRYNMQSPSKLQRLMKR---LPCGDRHCSPDHNGVOQVL 332
DB 310 --NLAAPSNCSLNPILISFLGETFRDKRLRYIEQKTNLPALDRFC--HAALRAVI 360

RESULT 15

US-09-970-820-38
Sequence 38, Application US/09970820
Patent No. US20020170077A1
GENERAL INFORMATION:

APPLICANT: FALB, DEAN A.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/970,820
FILING DATE: 05-Oct-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995

ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-032

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>

TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-970-820-38

Query Match 15.6%; Score 271; DB 9; Length 375;
Best Local Similarity 27.5%; Pred. No. 1.4e-16;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACQDLDGLSLSLGLVGVPGVCYNALVLANLHSAKSTMPDYFV 67
DB 44 NGTG--ELSEHQOYVIGL-FLSCLYTIFLPFIGVGNILILVNNISFREKKTIDLYFI 99
QY 68 NNAVAGLVLSALAPVHLLGPSSRWALMSVGGEVH--VALQIPF-NVSSIVAMYSTA 121
DB 100 NIAVADLLIVADSLIEVF-----NLHERYYDIAVLCTFMSLFLRVNMYSV 145
QY 122 L-----LSLDHYTERALPRTYMASVYNTRH---VCGFWGALITSFSSLLFYICSHYST 173
DB 146 FLITWMSFDRYI--ALARAMCSLFRTHHARLSGLIM--MASVSATLVPTAVHLOH 200
QY 174 RALECAKQNAEADATLVFIGYVP--ALATVVALVLSRVREDTFLDRDTGRLEP-- 229
DB 201 TDEACFCPADVREVOVMEVTLGFIIVFPALIGLCYSLIVRVLRH-----RHKG-LRPBR 254
QY 230 -SAHRLVAVTCTOGMTTPHYLLIGHTVLIISRGK-----VDANYLGHLHFVKD 279
DB 255 OKALRMIIAVALVVFVFCWLPENVFISVH--LLQRTQGAAPCKOSFRHAPL-TGHIV-- 309
QY 260 FSKLIAFSSPYTPPLVRYNMQSPSKLQRLMKR---LPCGDRHCSPDHNGVOQVL 332
DB 310 --NLAAPSNCSLNPILISFLGETFRDKRLRYIEQKTNLPALDRFC--HAALRAVI 360

Search completed: December 14, 2004, 20:27:30
Job time: 47.6006 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:04:39 ; Search time 14.5627 Seconds
(without alignments)
2200.157 Million cell updates/sec

Title: US-09-995-225B-2

Perfect score: 1733

Sequence: 1 MNCSEWNGGLVEELPAQO.....LPGDRHCHSPDHNGVQVLA 333

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267	15.4	375	2 JCS069	G protein-coupled
2	240.5	13.9	358	2 G02670	IL8-related recept
3	233.5	13.5	375	2 JCS509	G protein-coupled
4	205	11.8	358	2 A53752	interleukin-8 rece
5	200	11.5	360	2 A53671	interleukin-8 rece
6	193.5	11.2	350	2 A39445	interleukin-8 rece
7	185	10.7	355	2 JQ1231	interleukin-8 rece
8	184.5	10.6	369	2 JCS068	G protein-coupled
9	178.5	10.3	378	2 A55735	lymphocyte-specifi
10	177.5	10.2	378	2 B55735	G protein-coupled
11	177	10.2	372	2 S26657	interleukin-8 rece
12	175.5	10.1	367	2 JCS049	interleukin-8 rece
13	169	9.8	399	2 S29480	bombesin receptor
14	167	9.6	327	2 S56162	MDCK15 protein - h
15	167	9.6	354	2 I58186	probable G protein
16	166.5	9.6	352	2 I38973	G protein-coupled
17	166.5	9.6	352	2 A45747	neuropeptide Y/pep
18	166	9.6	359	2 A48921	interleukin-8 rece
19	165.5	9.5	378	2 A45680	G protein-coupled
20	165.5	9.5	388	2 JN0605	somatostatin recep
21	165.5	9.5	423	2 JCT677	allatostatin recep
22	164.5	9.5	353	2 S28787	neuropeptide Y/pep
23	164.5	9.5	359	2 A42656	angiotensin II rec
24	163	9.4	352	2 JCS0296	thyrotropin releas
25	162.5	9.4	384	2 A47249	brain-specific som
26	162	9.3	374	2 S42628	G protein-coupled
27	161.5	9.3	352	2 G00048	fusin (LESTRA) - c
28	161.5	9.3	354	2 A23659	interleukin-8 rece
29	161.5	9.3	356	2 S42096	interleukin-8 rece

30	161.5	9.3	359	2 JQ1516	angiotensin II rec
31	161.5	9.3	418	2 A46226	somatostatin recep
32	161	9.3	374	2 S32785	G protein-coupled
33	161	9.3	399	2 A46632	bombesin-like pept
34	160.5	9.3	366	2 S71152	neuropeptide Y/pep
35	160	9.2	355	2 A55733	G protein-coupled
36	159.5	9.2	391	2 A39297	somatostatin recep
37	158.5	9.1	363	2 I57940	somatostatin recep
38	158.5	9.1	391	2 C41795	somatostatin recep
39	158.5	9.1	428	2 S30508	probable G protein
40	157.5	9.1	428	2 A44021	somatostatin recep
41	155.5	9.0	384	2 JCS4629	somatostatin recep
42	155.5	9.0	391	2 A41795	somatostatin recep
43	155	8.9	362	2 I38990	Mel1b-melatonin re
44	155	8.9	362	2 A30341	G protein-coupled
45	154.5	8.9	350	2 JN0621	G protein-coupled

ALIGNMENTS

RESULT 1

JCS069

G protein-coupled receptor CMKRL2 - human

N:Alternate names: constitutively expressed peptide-like receptor, flow-induced endothe

C:Species: Homo sapiens (man)

C>Date: 31-Jan-1997 #sequence revision 28-Aug-1998 #text_change 09-Jul-2004

C/Accession: JCS069; JCS786; JCS294

R:Owman, C.; Blay, P.; Nilsson, C.; Lolait, S.J.

Biochem. Biophys. Res. Commun. 228, 285-292, 1996

A>Title: Cloning of human CDNA encoding a novel heptahelix receptor expressed in Burkitt

A:Reference number: JCS069; MVID:97079175; PMID:8920907

A/Accession: JCS069

A:Molecule type: mRNA

A:Residues: 1-375 <OM>

A/Cross-references: UNIPROT:Q99527; GB:Y08162; NID:g1707499; PIDN:CA69354.1; PID:g1707

A/Experimental source: B-cell lymphoblast

R:Tokada, Y.; Kato, C.; Kondo, S.; Korenaga, R.; Ando, J.

Biochem. Biophys. Res. Commun. 240, 737-741, 1997

A>Title: Cloning of cDNAs encoding G protein-coupled receptor expressed in human endoth

A:Reference number: JCS786; MVID:98063308; PMID:9398636

A/Accession: JCS786

A:Molecule type: mRNA

A:Residues: 1-375 <TK>

A/Cross-references: GB:AF015257; NID:g2353152; PIDN:AAC51904.1; PID:g2353153

A/Experimental source: umbilical vein endothelial cell

R:Feng, Y.; Gregor, P.

Biochem. Biophys. Res. Commun. 231, 651-654, 1997

A>Title: Cloning of a novel member of the G protein-coupled receptor family related to

A:Reference number: JCS294; MVID:97224403; PMID:9070864

A/Accession: JCS294

A:Molecule type: DNA

A:Residues: 1-311, 'T', 313-375 <FEN>

A/Cross-references: GB:U77827; NID:g1906591; PIDN:AAC51173.1; PID:g1906592

C/Comment: This protein plays a role in B-cell functions and is involved in endothelial

C/Genetics:

A:Gene: GDB:CMKRL2; CEPR

A/Cross-references: GDB:3929190

C/Superfamily: vertebrate rhodopsin

C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F:64-83/Domain: transmembrane #status predicted <TM1>

F:96-114/Domain: transmembrane #status predicted <TM2>

F:126-148/Domain: transmembrane #status predicted <TM3>

F:160-198/Domain: transmembrane #status predicted <TM4>

F:221-338/Domain: transmembrane #status predicted <TM5>

F:260-283/Domain: transmembrane #status predicted <TM6>

F:308-327/Domain: transmembrane #status predicted <TM7>

F:32,32,44/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 15.4%; Score 267; DB 2; Length 375;

Best Local Similarity 27.4%; Pred. No. 1,3e-14;

Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;

```

OY      8 NGTGLVEBLLPACODLOGLSLLSLGLVGVPGVLCYALLVLVLANHSKASMTMDYFV 67
Db      44 NGTG---ELSHQOYVIGL-FLSCLYTFLFPIGVGNILLVNINSFREQMTIPDLYFI 99
OY      68 MNAVGLVSLAPRHLLGPPSSRWALMSVQGEVHVALLQIFNVISLVAMSTYL-----L 123
Db      100 NLAVADLLVAVDSLIEVNFNHERYYDIAVLCTFMSFLQ-----VNMSSVFPFLTM 151
OY      124 SLDHYIERALPRTYMAVSYNTRH-----VCGFVMGALLTSPSSLFYCSHVSRALECA 179
Db      152 SFDKYI-ALBARARCSLFRKTHARKLSGLIM---MASVATLVPRVAHLQHTDEACF 206
OY      180 KMQNAEADATLVFPGVVP--ALATVAVLVLSVRREDPTDRDGTGLEP---SAHRL 234
Db      207 CFADREVQVMEVLTLGFIVPFALIGCYSLIVRVLVANH-----RHNG-LRPRQKALRM 260
OY      235 LVAVYCTOPGLMTPHYILILGHTVYIISRKP-----VDAYLGLLHFVKDFSKLLA 285
Db      261 ILAVLVFVQVMELEPNFISVH--LLORTQGAAPCKOSFPHAPL-TGHTIV---NLAA 313
OY      286 FSSSHVPLPLRYNMQNSPFSKLQRLMK---LPGDRHSCPDHMGVQOVL 332
Db      314 FSNCLNDLIVSLFGETFRDKRLRYISQKTNLPAINRC---NAALRAVI 360

```

RESULT 2
G02670
IL8-related receptor - human
C/Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C/Accession: G02670
R/McCoy, R.L., Perlmuter, D.H.
submitted to the EMBL Data Library, May 1996
A/Reference number: H01575
A/Accession: G02670
A/Status: preliminary, translated from GR/EMBL/DBD
A/Molecule type: mRNA
A/Residues: 1-358 <MCC>
A/Cross-references: UNIPROT:Q99527, EMBL:U58628, NID:gl381668, PIDN:AAE02736.1, PID:gl338
C/Genetics:
A/Gene: DRY12
A/Superfamily: vertebrate rhodopsin

Query Match	13.9%	Score 240.5;	DB 2;	Length 358;
Best Local Similarity	27.2%	Pred. No. 2e-12;		
Matches	90;	Conservative 61;	Mismatches 131;	Indels 49; Gaps 16;
QY	ELPACQDLQGLSLSLSLGLVGVVPGVCYALLVLNLNLSKASMTWPDVYFVVMAYAGL	74		
DB	ELSHQOQYIGL-FLSCYITTFLEFPGVGNILILVNISREKMTIPDLFYINLAIGDL	88		
QY	75 VLSALAPVHLIGPPSSRWALMSVGGEVHVALQIPFNVSILVAMYSTAL--LSLDHYTERA	132		
DB	89 ILVADSLIEVFNHERYYDIAVLCTFMSFLQVQH-----VQAASSFITMMSPDRIY--A	141		
QY	133 LPRYYMASVYNTRK----VCGFWGGLNLNFSLSL-FYICSHYSTALBCKAKQNNDEA	187		
DB	142 LARMRCSLFRTKHNARLSALIN-----MSSVSATLVPFTAVHLOHTDEACFCPADREV	197		
QY	188 DATLVFICYVVP--ALATLVALVLLSRVREDTFLDSDTRLEP---SAHRILVATVCTQ	242		
DB	198 QMDEVLTGFIIVFPAIITGLCYSLIYRVLVRAH-----RHRG-LRPBRKALRMILAVLVF	251		
QY	243 FGLMTFPHYLLLGHTVILISRGK-----VDAYILGLHFPYKPSKLLAESSFVTP	293		
DB	252 FVCLPENVFISVH--LLQRTQGAAPCKOSFRANHL-IGHIV---NLDAFNSGSLTP	304		
QY	294 LIRYNNQSPFSKQLRMKK---LPCGDRIIC	321		
DB	305 LIYSFLGETFRDKRLIYIEOKTNLPALNRF	335		

QCS509
 G:protein-coupled receptor 41 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 09-Jul-2004
 C:Accession: JCS509
 R:Bonini, U.A.; Anderson, S.M.; Steiner, D.F.
 Biochem. Biophys. Res. Commun. 234, 190-193, 1997
 A:Title: Molecular cloning and tissue expression of a novel orphan G protein-coupled receptor
 A:Reference number: JCS509; MUID:97312546; PMID:918987
 A:Accession: JCS509
 A:Molecule type: mRNA
 A:Residues: 1-375 <BON>
 A:Cross-references: UNIPROT:O08878; GB:U92802; NID:g2138341; PID:g213833
 A:Experimental source: lung
 C:Comment: This protein transduces the signal of a wide variety of hormones, neurotransmitters
 target cells.
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: glycoprotein, phosphoprotein
 F:32, 44, 84/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:86, 297/Binding site: phosphate (Ser) (covalent) #status predicted
 F:355/Binding site: phosphate (Thr) (covalent) #status predicted

	mac	67	consecutive	67	mismatch	138	inverts	32	days	16
OY	12	LVEELPACQDLOQLGSLLSLGLVGVPGVCYNALLVLANIHSKASMTWPDVYFVNMAV	71	:	:	:	:	:	:	:
Dd	45	LTGDSLHQOYVAL-FLSCLYTIFLEPPGFVNINILLVNISFPREKITPIDLYFINLAA	103	:	:	:	:	:	:	:
OY	72	AGVLISLAAPHLLGPSSRWALMSVGGEHVAILOIPEVNSLVAMYSTL-----LSLH	127	:	:	:	:	:	:	:
Dd	104	ADLLVLVADSLLEVFNDEQYYDDIAVLCTFMSLLOI-----NMYSVFPLTWMSFDR	155	:	:	:	:	:	:	:
OY	128	YIERALPRITMASVYNTNRH---VCGFWGALLTSFSSLFYICSHVSTRALECAKMON	183	:	:	:	:	:	:	:
Dd	156	YL--ALAKARCGLFRTKHARLSCGLIW---MASVASITLPPTVAHLRTEBACCPCPAD	210	:	:	:	:	:	:	:
OY	184	AEAADALVPIGYVP--ALATLVALVLSRVREDTPLDRODGRLEP---SHRLVLAT	238	:	:	:	:	:	:	:
Dd	211	VREVOMLEVTLGFIVPRPAIIIGLCYSILVKRLIRAH-----RHGG-LRPBRQKLXMIFAV	264	:	:	:	:	:	:	:
OY	239	VCTQFGMLTHYILLGHTVIISRKPVD-----HYGLGL-HEVKDFSKLLAESSSE	290	:	:	:	:	:	:	:
Dd	265	VLVFFICMLEBNPVISH-V-LIQMQOGDPTRCKOSFRHAPVPLGHIV----NLAAINSNG	318	:	:	:	:	:	:	:
OY	291	VTPLLRYRNQSPFSKJQLRMKK---LPCDRHCSDPHNGVOOYL	332	:	:	:	:	:	:	:
Dd	319	ISPILVIFDELFTPRDKLRVLVAOKTSLPALNRC--HATLKAVI	360	:	:	:	:	:	:	:

```

RESULT 4
A:53752
interleukin-8 receptor (clone 5B1a) - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #ext_change 09-Jul-2004
C:Accession: A53752
R:Prido, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.; Wilkinson, N.; Folco, E.; Navarric
J. Biol. Chem. 269, 12391-12399, 1994
A>Title: Molecular characterization of a novel rabbit interleukin-8 receptor isoform.
A:Reference number: A53752; MUID:94230294; PMID:8175642
A:Accession: A53752
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-358 <PPA>
A:Cross-references: UNIPROT:P53344; GB:L24445; NID:g437661; PIDN:AAA31378.1; PID:g437662
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match      11.8%; Score 205; DB 2; Length 358;
Best Local Similarity 23.0%; Pred. No. 1,8e-09;
Matches 70; Conservative 58; Mismatches 122; Indels 54; Gaps 9;

```



```

QY      1  MMSCSWFNGTGLVEELPACDLO-----LGLSLSLGLVGVVGL 42
      9  MMDDDLAFITGMP---PADEDYSPCMLETFETLNKRVYIAYALVFLSLSLG-----56
Db
QY      43  CYNALLVLANIHSKASMTMPDYVFVNNAVAGLVLSALAPVHLGPPSSRMALMSVGEVH 102
      57  --NSLVMLVILIXSVGSVTDVYLNLNLADLPLPALPLPI-----MAASRVNGWIF 105
Db
QY      103  VALQIPFNVSSTL---VAMSTALLSLDHYIERALPRTYMASVYNT-----RHVCGFV-- 151
      106  GTFLC--KVVSILKEVNFYSGIILLACISYDR-----YLAIVHATRTLQGRHLVKEVCL 158
Db
QY      152  --MGALLTSFSSLFFYCISVSTRALCAKQNAEADATVF-----IGVYVPLAT 203
      159  GCMGLSNMLSPFLFQAHYHNNSSPVCEVIGNDPAKRMVRLIPLHTFGFIVPLEFM 218
Db
QY      204  LYALVILSRVREDPTLDRDTGRLEPSAHRLVAVTCTQFGLMTPHYLILGHVITISR- 262
      219  LFCYGFILR-----TLFKAHMGQ-KHRAMVIFAVVLIPLLCMLPNVLVLLADTLMTQY 272
Db
QY      263  -----GKVDNHYIGLHFVKDFSKLAFSSSFVTPLLYRYMNSF 303
      273  IQSBCERRNNIGRALDA-----TEILGFHLSCINPIIYAFIQNF 312
Db

```

RESULT 7

```

J01231
interleukin-8 receptor - rabbit
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C/Accession: J01231; A46483
Biochem. Biophys. Res. Commun. 179, 784-789, 1991
R/Beckmann, M.P.; Munger, W.E.; Kozlowsky, C.; Vandenberg, T.; Price, V.; Lyman, S.; Getaz
A/Title: Molecular characterization of the interleukin-8 receptor.
A/Reference number: J01231; MUID:91378994; PMID:1898400
A/Accession: J01231
A/Molecule type: DNA
A/Residues: 1-355 <BEC>
A/Cross-references: UNIPROT:P21109; GB:M74240; NID:g165438; PIDN:AAA1375.1; PID:g165439
R/Lee, J.; Kang, W.J.; Rice, G.C.; Wood, W.I.
J. Immunol. 148, 1261-1264, 1992
A/Title: Characterization of complementary DNA clones encoding the rabbit IL-8 receptor.
A/Reference number: A46483; MUID:92148149; PMID:1737938
A/Accession: A46483
A/Status: Preliminary
A/Molecule type: mRNA
A/Residues: 1-355 <LEE>
A/Cross-references: GB:M82873; NID:g165440; PIDN:AAA1376.1; PID:g165441
A/Experimental source: neutrophils
A/Note: sequence extracted from NCBI backbone (NCBIN:81526, NCBI:P:81530)
C/Superfamily: vertebrate rhodopsin
C/Keyword: G protein-coupled receptor, transmembrane protein

```

Query Match 10.7%; Score 185; DB 2; Length 355;

Best Local Similarity 21.9%; Pred. No. 8e-08;

Matches 77; Conservative 65; Mismatches 133; Indels 76; Gaps 14;

```

QY      1  MMSCSWF-----NGTGL--VEE-----LPACDLO-----LGLSLSLGLVGVV 40
      11  LM--TFEDEFANATMPPEKDYSPCLVVTQTLNKKRVVYIAYALVFLSLSLG-----61
Db
QY      41  GLCYNALVLANIHSKASMTMPDYVFVNNAVAGLVLSALAPVHLGPPSSRMAL-----W 95
      62  ----NSLVMLVILIXSVGSVTDVYLNLNLADLPLPALMP1-----MAVSKEKGM 108
Db
QY      96  SVGEVHVALLQIPFNVSSTLVAMYSTALLSLDHYIERA-LPRTYMASVYNTRHVCGFVWG 154
      109  IGTPLCKKVVSLVKEVNFYSGIILLACISYDRYLAIVHATRTLQGRHLVKEFICIGIAL 168
Db
QY      155  ALLTSFSSLFFYCISVSTRALCAKQNAEADATVF-----IGVYVPLATLYALV 208
      169  SLTSLSPFLFQVPSVNNSSPVCEYEDLGHNTAKRMVRLIPLHTFGFIVPLVLMFCYG 228
Db

```

```

QY      209  LLSRVREDPTLDRDTGRLEPSAHRLVAVTCTQFGLMTPHYLILGHVIT-----I 260
      229  FTLR-----TLFQAHMGQ-KHRAMVIFAVVLIPLLCMLPNVLVLLADTLMTHTVIGERC 282
Db
QY      261  SRKPKDPAHYILGHVDFSKLAFSSSFVTPLLYRYMNSFPSTLQSLM 311
      283  QRKNDIDR-----ALDATEILGFHLSCINPIIYAFIQNFRNGFLKML 325
Db

```

RESULT 8

G protein-coupled receptor CCR-13 - human

C/Species: Homo sapiens (man)

C/Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 21-Jul-2000

C/Accession: J05068

R/Zaballos, A.; Varona, R.; Gutierrez, J.; Lind, P.; Marquez, G.

Biochem. Biophys. Res. Commun. 227, 846-853, 1996

A/Title: Molecular cloning and RNA expression of two new human chemokine receptor-like 9

A/Reference number: J05068; MUID:97040707; PMID:8886020

A/Accession: J05068

A/Molecule type: DNA

A/Residues: 1-369 <ZAB>

A/Cross-references: EMBL:Z79784; NID:g1668737; PIDN:CA802144.1; PID:g1668738

C/Comment: This protein belongs to the family of alpha chemokine receptors.

C/Genetic: GDB:CMKBR6; STBL22; GPR29; CCR6; CCR-13; GPR-CY4

A/Gene: GDB:CMKBR6; STBL22; GPR29; CCR6; CCR-13; GPR-CY4

A/Cross-references: GDB:5370639; OMIM:601835

A/Map position: 6q27-6q27

C/Superfamily: vertebrate rhodopsin

C/Keyword: G protein-coupled receptor, transmembrane protein

F/42-68/Domain: transmembrane #status predicted <TM1>

F/115-99/Domain: transmembrane #status predicted <TM2>

F/160-180/Domain: transmembrane #status predicted <TM3>

F/212-233/Domain: transmembrane #status predicted <TM4>

F/250-271/Domain: transmembrane #status predicted <TM5>

F/292-315/Domain: transmembrane #status predicted <TM6>

Query Match 10.6%; Score 184.5; DB 2; Length 369;

Best Local Similarity 21.1%; Pred. No. 9.1e-08;

Matches 68; Conservative 59; Mismatches 104; Indels 91; Gaps 11;

```

QY      41  GLCYNALVLANIHSKASMTMPDYVFVNNAVAGLVLSALAPVHLGPPSSRMALMSVGE 100
      55  GLGSLIVLVITFAFYKARSMTDVYLLMAADILFVLTLPFWVSHATGAMVFSNATCK 114
Db
QY      101  VHAL-QIPFNVSSTLVAMYSTALLSLDHYI-----ERALPRTYMASVYNTRHV 147
      115  LKGIYVAINFNCML-----LTCISMRYIAIVQATKSFRLRSKTLPR-----SKIT 162
Db
QY      148  GGFVWGALLTSFSSLFFYCISVSTRALCAKQNAEADATVF-----AIVIAVAVL 189
      163  CLVWGLSVIISSTFVF-----NOKYNTGSDVCEPKYQVSEPIRWKLM 209
Db
QY      190  ---TLVFGYVVPALATL--YALVLSRVREDPTLDRDTGRLEPSAHRLVAVTCTQFG 244
      210  LGLELRF-GFTPLPMFMTFCYTFIVKTLVQAKNSRKH-----AIVIAVAVL 260
Db
QY      245  LMTPHYLILGHVITISRKGEVDNHYIG-----LHFFVDFSKLAFSSSFVTP 294
      261  QQPHNNVLL-----VTANIGKNNRSCSEKLTGYTTVTEVLAFLHCLANPV 309
Db
QY      295  LYRYMNSFPSTLQSLM 316
      310  LYAFIQKFRNYFLKILKILWC 331
Db

```

RESULT 9

G protein-coupled receptor EB11 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004

Db 248 AGRPRQR-----KAVNALVTVSIFPLCMSPHIVFLDT--LAKLAVADNTCKLNGS 299
 QY 268 -----AHYGLHFVKDFSKLIAFSSSFVTPLLYRYMNSPFSKLORLMKKLP 316
 Db 300 LPVAITMCEPGLAH-----CCINPMLYTFAGVKFRSDLSRLTLTKGC 342

RESULT 12

JE0349
 Interferon-inducible protein 10 (IP-10) receptor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
 C/Accession: JE0349
 R/Tamaru, M.; Tomimaga, Y.; Yatsunami, K.; Narumi, S.
 Biochem. Biophys. Res. Commun. 251, 41-48, 1998
 A/Title: Cloning of the murine interferon-inducible protein 10 (IP-10) receptor and its
 A/Reference number: JE0349; MUID:9909219; PMID:9790904
 A/Accession: JE0349
 A/Molecule type: mRNA
 A/Residues: 1-367 <TMM>
 A/Cross-references: UNIPROT:Q9QWN6; DBU:AB003174; NID:g3798731; PIDN:BA434045.1; PID:g3
 C/Comment: This protein is important for lymphocyte trafficking to lymphoid organs.
 C/Superfamily: vertebrate rhodopsin

Query Match 10.1%; Score 175.5; DB 2; Length 367;
 Best Local Similarity 22.1%; Pred. No. 5,1e-07;
 Matches 77; Conservative 59; Mismatches 137; Indels 75; Gaps 14;

QY 8 NGGVLVEELPACODLQGL-----SLSLGLVGVPGVGLCYNALVLANLHNSKASMTM 61
 Db 32 NESDFSPSPCPDPSINFRTFLPALYSLLFL-----LGLNGVAVALLSQRTALSS 86
 QY 62 PDVYFVMAAGVLSALAPVHLGLPSSRWAL-----WSGGEVVALQIPNVSSLYA 116
 Db 87 TDFFLHLAVADVLLVLTPL-----WAVDAVQVWFGGLCKVAGALENINFYAG 137
 QY 117 MYSTALLSDHY-----IERALPRTYMASVYNTRHVCGFVWGALETSFSSLLFYIC 168
 Db 138 AFLACISFDRYLSIYHATQIYRDRVRVAL-----TCIYVWGCLFLALPDLF-LYS 190
 QY 169 SHVSTR--ALECAKMNAAADTVFI-----GYVVPALATY-----ALVLSRVRR 215
 Db 191 ANYDQRLNATHC--QYNFPQVGTALRVQLVAGFLPLPLVMAVCYAHILAVLSR--- 245
 QY 216 EDTPLDRDTGRLEPSAHLILVATVCTQFGMLTPHYLLGHYV---IISRKGVDAAHL 271
 Db 246 -----GQRFRLMKLVVVVVAFAVCMTFPHLVLVLDIMDVGLARCGRESH-- 294
 QY 272 GLIHFKYDFSKLIAFSSSFVTPLLYRYMNSPFSKLORLMKLP 319
 Db 295 --VDVASKVYTSIGWGMHCINPLLYAFVGVKFRFGWMLFTRLGSDQ 340

RESULT 13

S29480
 bombesin receptor - guinea pig
 C/Species: Cavia porcellus (guinea pig)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C/Accession: S29480
 R/Gorbulov, V.; Akhundova, A.; Beecher, H.; Fahrenholz, F.
 submitted to the EMBL Data Library, July 1992
 A/Description: Molecular cloning of a new bombesin receptor subtype.
 A/Reference number: S29480
 A/Accession: S29480
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-399 <GOR>
 A/Cross-references: UNIPROT:P35371; EMBL:X67126; NID:g49545; PIDN:CAA47605.1; PID:g49546
 C/Superfamily: endothelin receptor B
 C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 9.8%; Score 169; DB 2; Length 399;

Best Local Similarity 22.7%; Pred. No. 1,9e-06;
 Matches 68; Conservative 69; Mismatches 119; Indels 44; Gaps 14;

QY 38 VPVGLCYNALVLANLHNSKASMTMPDVYFVNMAVAG-LVLSALAPVHLGLPSSRWALMS 96
 Db 58 ISVGIIGNAILIKVFPEKTSMQTVENITFTSTALDGLLLTCVPDATHYLAEGWLFGR 117
 QY 97 VGEVVALQIPNVSSLYAMYSTALLSDHY--IERALPRTYMASVYNTRHVCGFVWCG 154
 Db 118 ICGKVSFIRL--TSVGVSFTLITISADRYKAVVPLERQPSNALIKTKAKACIWM 174
 QY 155 ALLTSFSSLLFYICGVST-----RALECA-----KQNAEADATLVFTGYVP 199
 Db 175 SMIFALPEAIF--SNVHTLRDPNKNMTSEWCIFYVSEKLOEIHALLSFLVF--YIIP 229
 QY 200 --ALATVALVLSHVRREDT---PLDRDT--GRLEPSAHLILVATVCTQFGV---WTPH 249
 Db 230 LSIISVYISLI--ATIVKSTINITEBQSHARKVYESRKRIAKTVLVVALFALCMLPN 287
 QY 250 YIILGHVTVISRGKPVDAHYGLHF-VKDFSKLIAFSSSFVTPLLYRYMNSPFSKLO 308
 Db 288 HLLNLVHSPTRKAYBDSA-----IHFIYVTFISRVLAFSNSCVNPFALYWLSTKTFQXQFK 342

RESULT 14

S56162
 MDCR15 protein - human
 C/Species: Homo sapiens (man)
 C/Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 08-Oct-1999
 C/Accession: S56162
 R/Barella, L.; Loetscher, M.; Tobler, A.; Baggiolini, M.; Moser, B.
 Biochem. J. 309, 773-779, 1995
 A/Title: Sequence variation of a novel heptahelical leucocyte receptor through alternat
 A/Reference number: S56162; MUID:95366951; PMID:7639692
 A/Accession: S56162
 A/Status: preliminary; nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Cross-references: EMBL:X68829; NID:g840783; PIDN:CAA48723.1; PID:g840784
 C/Superfamily: vertebrate rhodopsin

Query Match 9.6%; Score 167; DB 2; Length 327;
 Best Local Similarity 23.8%; Pred. No. 2,3e-06;
 Matches 77; Conservative 56; Mismatches 117; Indels 74; Gaps 15;

QY 27 SLSLGLVGVPGVGLCYNALVLANLHNSKASMTMPDVYFVNMAVAGVLSALAPVHLGL 86
 Db 14 SLIFLGVIGNVLV-----LVILERH-RQTRSTETPLFLHLAVADLLVLFIP-PAVA 64
 QY 87 PPSRWALMSVGEVVAL-QIPNVSSLYAMYSTALLSDHYIERALPRTYMASVYNT 145
 Db 65 EGSVGVAGTFLCTVIALHKNVFCSSL--LACTIAVDRLV-----AIVAHVAAYR 114
 QY 146 H-----VCGFVWGALETSFSSLLFYICS--HYSTRALCAKMNAAE--ADATLV 192
 Db 115 HRLSLIHICGTIMLVGLALPELFAVYSGCHHNSLRCTFSGENADLHAWTSSR 174
 QY 193 FI-----GYVVPALATYALV-LISRVAREDTPLDRDTGRLEPSAHLILVATVCTQFGWLT 247
 Db 175 FLYVNAQFLPLMLVMGVCYGVVHRLQDARRPQR-----KAVRAVILVTSIFFLCMS 228
 QY 248 PHYIILGHVTVISRGKPVV-----AHYGLHFVNDVDFSKLIAFSSSFVT 292
 Db 229 PTHIVIFLDT--LAKLAVADNTCKLNGSLPVAITMCEPGLAH-----CCIN 273
 QY 293 PLLYRYMNSPFSKLORLMKKLP 316
 Db 274 PMLYTFAGVKFRSDLSRLTLTKGC 297

RESULT 15

158186
 probable G protein-coupled receptor - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: I58186
R:Herliem, J.K.; Barber, C.M.; Lynch, K.R.
Neurosci. Lett. 169, 85-89, 1994
A:Title: CDNA cloning of a G-protein-coupled receptor expressed in rat spinal cord and H
A:Reference number: I58186; MUID:94323113; PMID:8047298
A:Accession: I58186
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>
A:Ctraa-reference: UNIPROT:P35411; EMBL:U04808; NID:G2558635; PIDN:AB87093.1; PID:9439
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor

Query Match 9.64; Score 167; DB 2; Length 354;
Best Local Similarity 22.94; Pred. No. 2.5e-06;
Matches 73; Conservative 54; Mismatches 120; Indels 72; Gaps 14;
QY 41 GLCYNLLVLANHSKSMTPDYFVMAVAGVLSALAPVHLGPPSSRWALMSVGE 100
DB 46 GLVGNLLVVALTNSRKSISITDIYLNIALSDLLFVATLPF-----WTHYLISSHE 96
QY 101 -VHAL-----QIPNVSLVAMYSTALLSDHYIERALPTYMASVYNTRHV-----C 148
DB 97 GLINAMCKLTARFFGIFPGGIFRITVISIDRYLAIVL-----AANSKMRITVQHGVTISL 152
QY 149 GFWGALTSFSSLLFYICSHVSTRALC-----AKQNAEADATLVFIQY 196
DB 153 G-VMAAIIIVASPOFMF-----TKRKNECIGDYEVLAGRIWPVLRNSE-----VNILGF 201
QY 197 VVRPALTLVALVLSRVRRDTPLDRTGRLPESAHRLVATVCTQGLWTPHYLILGH 256
DB 202 VLPULLMSFCYFRIVRT-----LPSCNRKKARAILLIVVVFFLFWTPYINIVFLE 255
QY 257 TV-----IISRKKPVDAHYLGLHFVDFSKLLAFSSSPVTPLLXYRYMNSPSPKLORLM 311
DB 256 TLKFYNFPPSCGKMRD-----LRMALSVETVAFSHCINPFIYAFAGEKFRIRYLRHLY 309
QY 312 KK--LPCGDRHCSPDHMG 327
DB 310 NKCLAVLCG----RPVHAG 324

Search completed: December 14, 2004, 20:25:44
Job time : 15.5627 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 18:37:44 ; Search time 54.3673 Seconds
(without alignments)
3524.170 Million cell updates/sec

Title: US-09-995-225B-2

Perfect score: 1733

Sequence: 1 MMSCSFNGTGLVEELPACO.....LPGDRHCSFPHMGVQVLA 333

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: UniProt 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1733	100.0	333	2	096CH1
2	1687	97.3	333	2	08HXR4
3	1277	73.7	333	2	08BHR6
4	1276	73.6	333	2	099LE2
5	1272	73.4	333	2	08BHU4
6	1092	63.0	333	2	06PTG9
7	1092	63.0	333	2	AAH61674
8	772	44.5	151	2	086SP5
9	267	15.4	375	1	CML2_HUMAN
10	261	15.1	375	1	06FNU6
11	233.5	13.5	375	1	CML2_RAT
12	229.5	13.2	375	1	09D392
13	223.5	12.9	375	2	08BMP4
14	208.5	12.0	355	2	08HZN5
15	206.5	11.9	355	2	08HZN3
16	205	11.8	358	1	ILB8_RABIT
17	203.5	11.7	355	1	ILB8_RABIT
18	201.5	11.6	353	1	ILB8_MACMU
19	201	11.6	353	1	ILB8_PANTR
20	201	11.6	355	2	08HZN7
21	200	11.5	353	1	ILB8_GORCO
22	200	11.5	355	2	08HZN8
23	198.5	11.5	350	1	ILB8_HUMAN
24	198.5	11.5	350	1	ILB8_GORCO
25	198	11.4	352	2	0810T4
26	195.5	11.3	350	2	06IN95
27	195.5	11.3	374	2	08HXR7
28	193.5	11.2	350	1	ILB8_HUMAN
29	193.5	11.2	350	2	08N6T6
30	192.5	11.1	517	2	09VWR3
31	191	11.0	350	1	ILB8_PANTR

32	190	11.0	366	2	0867B2	0867B2 capra hircu
33	188	10.8	355	2	08HZN6	08hzn6 pongo pygma
34	185	10.7	349	1	ILB8_RAT	P70612 ratius norv
35	185	10.7	355	1	ILB8_RABIT	P21109 oryctolagus
36	184.5	10.6	374	1	CKRE_HUMAN	P51684 homo sapien
37	183	10.6	380	2	0861S1	0861s1 sus scrofa
38	183	10.6	380	2	BAD06309	BAD06309 sus scrofa
39	181.5	10.5	378	2	06U2D6	06u2d6 ratius norv
40	181.5	10.5	378	2	AAK24573	AAK24573 ratius no
41	180.5	10.4	359	2	09PYV7	09pyv7 anguilla an
42	180	10.4	359	1	CCR4_MOUSE	P70658 m-c-x-c che
43	179.5	10.4	378	2	08CA82	08ca82 mus musculu
44	179	10.3	351	2	0810W6	0810w6 mus musculu
45	179	10.3	351	2	AAQ84305	AAQ84305 mus muscu

ALIGNMENTS

RESULT 1	ID	Q96CH1	PRELIMINARY	PRT	333 AA
AC	Q96CH1				
DT	01-DEC-2001	(TrEMBLrel. 19, Created)			
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)			
DE	G protein-coupled receptor 146.				
GN	Name=GPR146;				
OS	Homo sapiens (human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSE=Kidney;				
RX	MDLIRN=22388257; PubMed=12477932;				
RA	Straube R.L., Peingold E.A., Grouse L.H., Derge J.G.,				
RA	Klaener R.D., Collins F.S., Wagner L., Shemen C.M., Schlier G.D.,				
RA	Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mulhany S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Pahey J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,				
RA	Krzywnicki M.I., Skalska U., Smaluk D.E., Schnerch A., Schein J.E.,				
RT	Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSE=Kidney;				
RA	Straube R.,				
RL	Submitted (SEP-2001) to the EMBL/GenBank/DOBJ databases.				
DR	EMBL, BC014241, AAH14241.1; -.				
DR	Genew, HENC:21718, GPR146.				
DR	GO, GO:0016021, C:integral to membrane, IEA.				
DR	GO, GO:0004872, F:receptor activity, IEA.				
DR	GO, GO:0001584, F:rhodopsin-like receptor activity, IEA.				
DR	GO, GO:0007186, P:G-protein coupled receptor protein signalin. . .; IEA.				
DR	InterPro, IPR000276; GPCR_Rhodopsin.				
DR	Pfam, PF00001; 7tm_1, 1.				
DR	PRINTS, PR00237; GPCRHHODOPSN.				
DR	PROSITE, PSS0262; G_PROTEIN_RECPT_FL_2, 1.				
KW	Receptor.				
SQ	SEQUENCE 333 AA; 36580 MW; BB8D6505A6C2D9F4 CRC64;				

Query Match 100.0%; Score 1733; DB 2; Length 333;
Best Local Similarity 100.0%; Pred. No. 1,6e-122;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCGWNGTGLVEELPACODLQIGLSLSLGLVGVPGVGCYNALLVLANLHKSASMT 60
DB 1 MMSCGWNGTGLVEELPACODLQIGLSLSLGLVGVPGVGCYNALLVLANLHKSASMT 60
QY 61 MPDYVFVNMAVAGLVISALAPVHLGPPSSRMALMSVGEVHALQIPNVSSLVAMYST 120
DB 61 MPDYVFVNMAVAGLVISALAPVHLGPPSSRMALMSVGEVHALQIPNVSSLVAMYST 120
QY 121 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 121 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
QY 181 MONEAADATLVFIGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
DB 181 MONEAADATLVFIGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
QY 241 TQFGLWTPHYLLILGHTVLIISRGKPYDAHYLGILHFVKDPSKLLAFSSSFVTPPLYRYNN 300
DB 241 TQFGLWTPHYLLILGHTVLIISRGKPYDAHYLGILHFVKDPSKLLAFSSSFVTPPLYRYNN 300
QY 301 QSFPSKLRMLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSFPSKLRMLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 2
Q8HXP4 PRELIMINARY; PRT; 333 AA.
ID Q8HXP4;
AC Q8HXP4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypochemical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal lobe left;
RX MEDLINE=21458551; PubMed=11574149;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirata M., Suto Y.,
RT Hirata M., Terao K., Suzuki Y., Sugano S., Hashimoto K.,
RT "Assignment of 118 novel cDNAs of cynomolgus monkey brain to human
RT chromosomes.";
RL Gene 275:31-37(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal lobe left;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Sugano S.,
RT Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB093636; BAC21610.1; -;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signaling. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00262; G_PROTEIN_RECPT_F1_2; 1.
KW Hypochemical protein.
SQ SEQUENCE 333 AA; 36365 MW; 9D955B0461F93439 CRC64;

Query Match 97.3%; Score 1687; DB 2; Length 333;
Best Local Similarity 96.4%; Pred. No. 4,6e-119;
Matches 321; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 MMSCGWNGTGLVEELPACODLQIGLSLSLGLVGVPGVGCYNALLVLANLHKSASMT 60

DB 1 MMSCGWNGTGLVEELPACODLQIGLSLSLGLVGVPGVGCYNALLVLANLHKSASMT 60
QY 61 MPDYVFVNMAVAGLVISALAPVHLGPPSSRMALMSVGEVHALQIPNVSSLVAMYST 120
DB 61 MPDYVFVNMAVAGLVISALAPVHLGPPSSRMALMSVGEVHALQIPNVSSLVAMYST 120
QY 121 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 121 ALLSLDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
QY 181 MONEAADATLVFIGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
DB 181 MONEAADATLVFIGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
QY 241 TQFGLWTPHYLLILGHTVLIISRGKPYDAHYLGILHFVKDPSKLLAFSSSFVTPPLYRYNN 300
DB 241 TQFGLWTPHYLLILGHTVLIISRGKPYDAHYLGILHFVKDPSKLLAFSSSFVTPPLYRYNN 300
QY 301 QSFPSKLRMLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSFPSKLRMLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 3
Q8HXR6 PRELIMINARY; PRT; 333 AA.
ID Q8HXR6;
AC Q8HXR6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length
DE enriched library, clone:B230399N13 product:HYPOHETICAL 36.6 kDa
DE PROTEIN, full insert sequence.
GN Name=B230399N13;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.,
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA The PANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=2049574; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;

RX MEDLINE=20530913; PubMed=11076861;
 RA Shiba K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama U., Nishi K., Kitahara T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuyama S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RP [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Corpora quadrigemina;
 RA Adecchi U., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kaikawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shiba K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takeku-Akahira S., Tanaka Y., Tanaka T.,
 RA Tomaru A., Taya T., Yamanishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL/AC046512; BAC32762.1; -
 DR MGD; MGI:1933113; BC003323.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE; PRO0237; GPCR_HODOPSIN.
 DR PRINTS; PS50262; G_PROTEIN_RECPR_F1_2; 1.
 DR KX Hypothetical protein.
 SQ SEQUENCE 333 AA; 36571 MW; CS1690838AB42988 CRC64;

Query Match 73.7%; Score 1277; DB 2; Length 333;
 Best Local Similarity 74.5%; Pred. No. 3.6e-88;
 Matches 248; Conservative 31; Mismatches 52; Indels 2; Gaps 2;

QY 1 MMSCSFNGTGLVVEELPACODLQLGLSLSLGLGVVGPVGLCYNALVLANHSKASMT 60
 DB 1 MMSCGPLNSTAMER-PLCNRLRLGLMVLSTLYLGAVPPLSLGNALVLANIASKNTMT 59
 QY 61 MPDVYEVNNAVAGLVSLAPVHLGLPSSSRMALWVGVEVHALQIPENVSSLVAMYST 120
 DB 60 MPDVYEVNNAVAGLVSLAPVHLGLPSSSRMALWVGVEVHALQIPENVSSLVAMYST 119
 QY 121 ALISLDHYERAPRTYMASVYTRHYCGFVWGALITSSSLFTYICSHVSTRALBCK 180
 DB 120 ALISLDHYERAPRTYMASVYTRHYCGFVWGALITSSSLFTYICSHVSTRALBCK 179
 QY 181 MONAABADTLVIGVVPVPLATLVLTALVLSRVRRBDPLDRDTGLBPAHLLVAATVC 240
 DB 180 MONAABADTLVIGVVPVPLATLVLTALVLSRVRRBDPLDRDTGLBPAHLLVAATVC 239
 QY 241 TOFGLTPTHTLLGLHTVLIISRGKPYDAHYLGLHFKVPSKLLAASSSFVPTLVRYNN 300
 DB 240 TOFGLTPTHTLLGLHTVLIISRGKPYDAHYLGLHFKVPSKLLAASSSFVPTLVRYNN 298
 QY 301 QSEPSKLORLMKGLPGCDRHCSPDHNGVQOYLA 333
 DB 299 KAFPGKLRILMKMKHCGRRHCSPDPSGICQVMA 331

RESULT 4
 Q99LE2 PRELIMINARY; PRT; 333 AA.
 AC Q99LE2;
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 01-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE CDNA sequence BC003333 (Mus musculus NOD-derived Cplc +ve dendritic
 DE cells cDNA, RIKEN full-length enriched library, clone:630505002
 DE product:HYPOPHYSICAL 36.6 Kda PROTEIN, full insert sequence).
 OS Name=BC003333;
 GN Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feilgold E.A., Grouse L.H., Derge J.G.,
 RA Klautner R.D., Collins F.S., Wagner L., Steinem C.M., Schler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.B.,
 RA Brownstein M.J., Udén T.B., Toshynski S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Petera G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN PANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [5]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RA The PANTOM Consortium;
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team:
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [6]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shiba K., Hayatsu N., Sugahara Y., Shiba K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [7]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:11757-1771(2000).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AK089323; BAC30804.1; -;
 DR MGD; MGI:1933113; BC003323.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm1, 1.
 DR PRINTS; PR00337; GPCRHOOPS.
 DR PROSITE; PS0262; G_PROTEIN_RECPT_FL_2; 1.
 DR Hypothetical protein.
 KW SEQUENCE 333 AA; 36557 MW; 5F76909350D801F4 CRC64;
 SQ
 Query Match 73.6%; Score 1276; DB 2; Length 333;
 Best Local Similarity 74.5%; Pred. No. 4.3e-88;
 Matches 248; Conservative 31; Mismatches 52; Indels 2; Gaps 2;
 QY 1 MMSGSGWNGTGLVEELPACODLQGLSLSLGLVVGVPVGCYNAVLTLANLHAKASMT 60
 DB 1 MMSGGLPSTNAMAEE-PLCGNRLGLGLWVSLILYLGAGVPSIGYNALLVLANLAKMTWT 59
 QY MPVYFPNNAVAGLVLSALAPVHLPPSSRWALMSVSGEVAVALQIPNVSLVAMST 120
 DB 60 MPVYFPNNAVAGLVLTALAPVHLPPSSRWALMSVSGEVAVALQIPNVSLVAMST 119
 QY 121 ALLSLDHYTERALPRTYMASVYTRHVCFFVWGGLTSPSSILTYITCSHSTRALBCAK 180
 DB 120 ALLSLDHYTERALPRTYMASVYTRHVCFFVWGGLTSPSSILTYITCSHSTRALBCAK 179
 QY 181 MONAARAADATVFIGVVPALATLVALLSLSVREEDPPLDRTGRLPSAHLRLVATVC 240
 DB 180 MONTEAADAALVAVIGVVPGLVALLVALISRIKEDPPLDQDTRSLDPSVHRLVATVC 239
 QY 241 TQFGMLTPHYLLILGHTVVISRGKVPDAHYGLHFPVDFSKLAFSSFTPLLYRYMN 300
 DB 240 TQFGMLTPHYLL-SLGHTVVISRGKRVTEGVEHGLIGVADLAKFLAFSSSVTPPLLYRYIN 298
 QY 301 QSFPSKQLQRLMKLPCGSRHSCPDHMGVQVYLA 333
 DB 299 KAFPGKLRRLMKMKMGCRHRSPPDSGIGQVMA 331
 RESULT 5
 OGBHU4 PRELIMINARY; PRT; 333 AA.
 AC OGBHU4;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched
 DE library, clone:R030026M12 product:HYPOTHEICAL 36.6 kDa PROTEIN, full
 DE insert sequence.
 GN Name=BC003323;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=99279253; PubMed=10349636;
 RT Carninci P., Hayashizaki Y.;
 RL "High-efficiency full-length cDNA cloning.";
 RM Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RL "Functional annotation of a full-length mouse cDNA collection.";
 RN Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RA The FANTOM Consortium;
 RL "The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RL "Normalization and subraction of cap-trapper-selected cDNAs to
 RL prepare full-length cDNA libraries for rapid discovery of new genes.";
 RN Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:11757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AK087100; BAC39803.1; -;
 DR MGD; MGI:1933113; BC003323.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.

DR GO:0007186; P:G-protein coupled receptor protein signalin. . . IEA.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm.1; 1
 DR PRINTS, PR00237; GPCRHOOPS.
 DR PROSITE, PSS0262; G_PROTEIN_RECEP_F1_2, 1.
 KW Hypothetical protein.
 SQ SEQUENCE 333 AA; 36597 MW; C5168D28B71F5988 CRC64;

Query Match 73.4%; Score 1272; DB 2; Length 333;
 Best Local Similarity 74.2%; Pred. No. 8.5e-86;
 Matches 247; Conservative 31; Mismatches 53; Indels 2; Gaps 2;

QY 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVPGVGLCYNALLVLANHKSAMT 60
 DB 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVPGVGLCYNALLVLANHKSAMT 59
 QY 61 MPDYFVNMAVAGLVLTALPAVLLGPAHSMALMSISEAHVTLILFVNASLVVYST 120
 DB 60 MPDYFVNMAVAGLVLTALPAVLLGPAHSMALMSISEAHVTLILFVNASLVVYST 119
 QY 121 ALLSLDHYIERALPRTYMASVYVTRHYCGFVWGALVTSFSSLLFYICSHVSTRALBCKA 180
 DB 120 ALLSLDHYIERALPRTYMASVYVTRHYCGFVWGALVTSFSSLLFYICSHVSTRALBCKA 179
 QY 181 MONAADAATLVFVIGVVPALATLYALVLSRVAREPTPLDRDTRGLEPSAHRLVATVC 240
 DB 180 MONAADAATLVFVIGVVPALATLYALVLSRVAREPTPLDRDTRGLEPSAHRLVATVC 239
 QY 241 TOFGLMTPHYLLLGHTVVISRGKPVDAHYLGILHFKVDSKLLAFSSFPVTPLLYRYN 300
 DB 240 TOFGLMTPHYLLLGHTVVISRGKPVDAHYLGILHFKVDSKLLAFSSFPVTPLLYRYN 298
 QY 301 QSFPSKLGRLMKLPCGDRHCSPDHMGVQVLA 333
 DB 299 KAFPGKLRMLKMKPCGRHCSPPSGIQVMA 331

RESULT 6

Q6P7G9 PRELIMINARY; PRT; 333 AA.

AC 06P7G9; (TEMBLrel. 27, Created)
 DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
 DE MGC68817 protein.
 GN Name=MGC68817;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OK NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=22368257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusik A., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullan S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Kravitz M.I., Skalkida U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT Initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]

RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Klein S., Strausberg R.;
 RU Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC061674; AAH61674.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS, PR00237; GPCRHOOPS.
 DR PROSITE, PSS0262; G_PROTEIN_RECEP_F1_2, 1.
 SQ SEQUENCE 333 AA; 38007 MW; EB72DC50ECBBE815 CRC64;

Query Match 63.0%; Score 1092; DB 2; Length 333;
 Best Local Similarity 60.8%; Pred. No. 3.1e-74;
 Matches 202; Conservative 61; Mismatches 69; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVPGVGLCYNALLVLANHKSAMT 60
 DB 1 MMSCEDLNTYNSGEBOYLCHNEFHLFLIFSVLLVILICFPVGLCYNQLVYVNLNKAHMT 60
 QY 61 MPDYFVNMAVAGLVLTALPAVLLGPAHSMALMSISEAHVTLILFVNASLVVYST 120
 DB 61 MPDYFVNMAVAGLVLTALPAVLLGPAHSMALMSISEAHVTLILFVNASLVVYST 120
 QY 121 ALLSLDHYIERALPRTYMASVYVTRHYCGFVWGALVTSFSSLLFYICSHVSTRALBCKA 180
 DB 121 TLISLDYIERALPRTYMASVYVTRHYCGFVWGALVTSFSSLLFYICSHVSTRALBCKA 180
 QY 181 MONAADAATLVFVIGVVPALATLYALVLSRVAREPTPLDRDTRGLEPSAHRLVATVC 240
 DB 181 MONAADAATLVFVIGVVPALATLYALVLSRVAREPTPLDRDTRGLEPSAHRLVATVC 240
 QY 241 TOFGLMTPHYLLLGHTVVISRGKPVDAHYLGILHFKVDSKLLAFSSFPVTPLLYRYN 300
 DB 241 TOFGLMTPHYLLLGHTVVISRGKPVDAHYLGILHFKVDSKLLAFSSFPVTPLLYRYN 300
 QY 301 QSFPSKLGRLMKLPCGDRHCSPDHMGVQVLA 332
 DB 301 KNFSGKLRMLKMKPCGRHCSPPSGIQVMA 332

RESULT 7

AAH61674 PRELIMINARY; PRT; 333 AA.

AC AAH61674;
 DT 02-MAR-2004 (TEMBLrel. 27, Created)
 DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TEMBLrel. 27, Last annotation update)
 DE MGC68817 protein.
 GN Name=MGC68817;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OK NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RT inactive."
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang U., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stampleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Yoshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lonnellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., Mcwan P.O., McKernan K.U., Malek U.A., Gunaratne P.H.,
RA Richard S., Morley K.C., Hale S., Garcia A.M., Gay L.U., Huiyik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywiński M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Klein S., Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC061674; AAH61674.1; -
SQ SEQUENCE 333 AA; 38007 MW; EB72DC50EC8EB15 CRC64;

Query Match 63.0%; Score 1092; DB 2; Length 333;
Best Local Similarity 60.8%; Pred. No. 3.1e-74;
Matches 202; Conservative 61; Mismatches 69; Indels 0; Gaps 0;

QY 1 MMSGSFNGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNMLLVIANHKSASMT 60
DB 1 MMSCEDLNTNGSEGYLNCNEFLFLFISVLYLIICFPYGLCYNQLVLYNKNATMT 60
QY 61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRWALMSVGEVHALQIPFNVSILVAMYS 120
DB 61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRWALMSVGEVHALQIPFNVSILVAMYS 120
QY 121 ALISLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALCA 180
DB 121 TLLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALCA 180
QY 181 MKNAEADATLVFIGYVVPALATLYALVLSRVREDTPLDRDTGRLPSAHGLVATYC 240
DB 181 MKNREAADATLVFIGYVVPALATLYALVLSRVREDTPLDRDTGRLPSAHGLVATYC 240
QY 241 TOFGLTTPHYLLDGHVTIISRGKPYDAHYLGHLFVKDFSKLAFSSSFVTPLYRYNM 300
DB 241 TOFGLTTPHYLLDGHVTIISRGKPYDAHYLGHLFVKDFSKLAFSSSFVTPLYRYNM 300
QY 301 QSFPSKLGRIKMKLLPCGDRHCSFDHNGVOQL 332
DB 301 KNFSGKLGRIKMKLLPCGDRHCSFDHNGVOQL 332

RESULT 8
Q86SP5 PRELIMINARY; PRT; 151 AA.
AC Q86SP5; (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE G protein-coupled receptor, FGR8 (Fragment).
OS Homo sapiens (Human).
RA Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22584407; PubMed=12679517;
RA Vassiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
RA Bergman J.E., Galtanaris G.A.;
RT "The G protein-coupled receptor repertoires of human and mouse."
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).
DR EMBL: AY255535; AA085047.1; -
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein-coupled receptor protein signalin. . .; IEA.
DR Interpro: IPR000276; GPCR_Rhodpsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KM Receptor.
FT NON_TER 1 151
FT NON_TER 151 151
SQ SEQUENCE 151 AA; 16325 MW; 42FE8329C47C699A CRC64;

Query Match 44.5%; Score 772; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8e-50;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 TMDYVFVNMAVAGLVLSALAPVHLGPPSSRWALMSVGEVHALQIPFNVSILVAMYS 119
DB 1 TMDYVFVNMAVAGLVLSALAPVHLGPPSSRWALMSVGEVHALQIPFNVSILVAMYS 60
QY 120 TALLSDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALCA 179
DB 61 TALLSDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALCA 120
QY 180 KKNAEADATLVFIGYVVPALATLYALVYL 210
DB 121 KKNAEADATLVFIGYVVPALATLYALVYL 151

RESULT 9
CML2_HUMAN STANDARD; PRT; 375 AA.
AC Q99537; O00143; Q43494; Q13631; Q96F42; Q99981;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chemokine receptor-like 2 (IL8-related receptor DRY12) (Flow-induced endohepial G protein-coupled receptor) (FEG-1) (G protein-coupled receptor GPR30) (GPCR-BR).
DE Name=GPR30; Synonyms=CMKRL2, DRY12, CEPR;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97079175; PubMed=8920907;
RA Owan C.S.O., Bly P., Nilsson C., Lolait S.J.;
RT "Cloning of human cDNA encoding a novel heparin-like receptor expressed in Burkitt's lymphoma and widely distributed in brain and peripheral tissues."
RL Biochem. Biophys. Res. Commun. 228:285-292 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97286314; PubMed=9141481;
RA Kvingedal A.M., Smeland E.B.;
RT "A novel putative G-protein-coupled receptor expressed in lung, heart and lymphoid tissue."
RL FEBS Lett. 407:59-62 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97224403; PubMed=9070864;
Feng Y., Gregor P.;

RT "Cloning of a novel member of the G protein-coupled receptor family
 RT related to peptide receptors.";
 RT Blochem. Biophys. Res. Commun. 231:651-654(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver,
 RA MCCOY R.L., Perlmuter D.H.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98063308; PubMed=9398636;
 RA Takahashi Y., Kato C., Kondo S., Korenaga R., Ando J.;
 RT "Cloning of cDNAs encoding G protein-coupled receptor expressed in
 RT human endothelial cells exposed to fluid shear stress.";
 RL Biochem. Biophys. Res. Commun. 240:737-741(1997).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98035885; PubMed=9367686;
 RA Carmeci C., Thompson D.A., Ring H.Z., Francke U., Weigel R.J.;
 RT "Identification of a gene (GPR30) with homology to the G-protein-
 RT coupled receptor superfamily associated with estrogen receptor
 RT expression in breast cancer.";
 RL Genomics 45:607-617(1997).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98140132; PubMed=9479505;
 RA O'dowd B.P., Nguyen T., Marchese A., Cheng R., Lynch K.R.,
 RA Heng H.H.Q., Kojakowski L.F. Jr., George S.R.;
 RT "Discovery of three novel G-protein-coupled receptor genes.";
 RL Genomics 47:310-313(1998).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Uterus;
 RX MEDLINE=92388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares W.B., Bonaldo M.F., Casavant T.U., Scheetz T.E.,
 RA Brownstein M.J., Uebachs T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosack S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smalua D.E.,
 RA Scheraga A., Schein J.E., Jones S.J.W., Maria M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Orphan receptor, possibly for a chemokine.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in lung, heart, liver, brain and
 CC lymphoid tissue. In brain regions, expressed as a 2.8 kb
 CC transcript in basal forebrain, frontal cortex, thalamus,
 CC hippocampus, caudate and putamen.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; Y08162; CAA69354.1; -
 DR EMBL; X98510; CAA67133.1; -
 DR EMBL; U77827; AAC51173.1; -

DR EMBL; U58828; AAB02736.1; -
 DR EMBL; AF015257; AAC1904.1; -
 DR EMBL; U63917; AAB8017.1; -
 DR EMBL; AF027956; AAC52027.1; -
 DR EMBL; BC01634; AAH1634.1; -
 DR PIR; G02670; G02670.
 DR PIR; JCS069; JCS069.
 DR Genew; HGNC:4485; GPR30.
 DR MIM; 601805; -
 DR GO; GO:0005887; C: integral to plasma membrane, TAS.
 DR GO; GO:0004930; F: G-protein coupled receptor activity, TAS.
 DR GO; GO:0007186; P: G-protein coupled receptor protein signaling. . . ; TAS.
 DR InterPro; IPR00276; GPCR_Rhodopsin.
 DR InterPro; IPR002286; P2_purinocptor.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsin.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1;
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 FT DOMAIN 1 62
 FT TRANSMEM 63 84
 FT DOMAIN 85 96
 FT TRANSMEM 97 120
 FT DOMAIN 121 132
 FT TRANSMEM 133 153
 FT DOMAIN 154 175
 FT TRANSMEM 176 194
 FT DOMAIN 195 220
 FT TRANSMEM 221 236
 FT DOMAIN 237 259
 FT TRANSMEM 260 280
 FT DOMAIN 281 306
 FT TRANSMEM 307 327
 FT DOMAIN 328 375
 FT DISULFID 130 207
 FT CARBOHYD 25 25
 FT CARBOHYD 32 32
 FT CARBOHYD 44 44
 FT TRANSMEM 16 21
 FT TRANSMEM 20 21
 FT TRANSMEM 32 32
 FT TRANSMEM 104 104
 FT TRANSMEM 140 146
 FT TRANSMEM 179 179
 FT TRANSMEM 184 184
 FT TRANSMEM 312 312
 FT TRANSMEM 320 320
 FT TRANSMEM 355 355
 FT TRANSMEM 358 358
 SQ SEQUENCE 375 AA, 42247 MW, 0A54EBDA658F075 CRC64;
 Query Match 15.4%; Score 267; DB 1; Length 375;
 Best local similarity 27.4%; Pred. No. 5.1e-12;
 Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;
 QY 8 NGGVIVELPACQDLQGLSLSLGLGVNVPVGLCNALVLANLSKSKSMTPDQVYF 67
 DB 44 NGGV--ELSEHQVYIGL-FLSCLTYTFPLPGVGNILIVNVSFRKMTIPDLVFI 99
 QY 68 NMAVAGVLSALAPVHLGPPSSRWALMSVGEVHVLAQIPFNVSILVAMYSTAL--L 123
 DB 100 NLAVADILVADSLIEVFNHERVYDVAVCTMSLFLQ-----VNMYSSVFLTWM 151
 QY 124 SLDIYERLPRVYMASVYVTRH---VCGFWGALLISFSSLLFYTCHSVSTRALCA 179
 DB 152 SPDRYI--ALARMSCSLFRTKHARLSGLIW---MASVSATVVFPTAVHLQHTBEACF 206
 QY 180 KMGNAAADATVIFIGVVP--ALATLYALVLSRYRRETPVDRPGRLEP---SAHRL 234
 DB 207 CPADVREVQWLEVTGLGIVPAITIGCYSLIVAVLVAH---RRHG-LRPRRQALRM 260
 QY 235 LVAVTCTQFGLMTPHYLLGLHTVILSRGR-----VDARYGLGLHFVDFSLTALA 285

```

Db      261 ILAVLVFVFCMLPENVFISVH--LLQRTQPGACAPCKOSFRHAHL-TGHIV---NLAA 313
OY      286 FSSSFPTPLLYRYMNSPFSKQRLMK--LPCGDRHCSPDHMGVQVLT 332
Db      314 FNSNCINPLIYSFLGTFPRDKRLTYEQKTNLPALNRF--HAALKAVI 360

RESULT 10
O6FHU6  PRELIMINARY; PRT; 375 AA.
ID      06FHU6
AC      06FHU6;
DT      05-JUL-2004 (TREMBLrel. 27, Created)
DT      05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE      05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE      GPR30 Protein (Fragment).
GN      Name=GPR30;
OS      Homo sapiens (Human)
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Hallack A., Ebert L., Moundinya M., Schick M., Eisenstein S.,
RA      Neubert P., Ketrang K., Schalten R., Shen B., Henze S., Mar W.,
RA      Korn B., Zhu D., Hu Y., Labaer J.;
RL      Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC      -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
DR      EMBL; CR541655; CAC646456.1;
DR      InterPro; IPR000276; GPCR_Rhodopsn.
DR      InterPro; IPR002286; P2_purinocptor.
DR      Pfam; PF00001; 7tm_1; 1.
DR      PRINTS; PR00237; GPCR_RHODOPSIN.
DR      PRINTS; PR01157; P2PYRNOCPTR.
DR      PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR      PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW      G-protein coupled receptor; Receptor; Transmembrane.
FT      NON TER 375
SQ      SEQUENCE 375 AA; 42194 MW; 5A54FDBA2E36AE CRC64;

Query Match 15.1%; Score 261; DB 2; Length 375;
Best Local Similarity 26.9%; Pred. No. 1.5e-11;
Matches 94; Conservative 66; Mismatches 132; Indels 58; Gaps 17;

OY      8 NGTGLVBEIPACDDLGSLSLGLGVGVPGCTYNALVLANLHASKASMTMPDVYFV 67
Db      44 NGTG---ELSEHQGYVIGL-PLSCLYTIPLFPIGFVGNILIVVNISFREKKTIPDLVFI 99
OY      68 NMAVAGLVLSALPVLHLPSSSRMALMSVGEVHALQIPFVNSLVAMYSTAL---L 123
Db      100 NLAVADLILVADSLIEVFNHRYVIAVLCTFMSLFLQ-----VNMYSVFELTWM 151
OY      124 SLDHYIERALPRTYMASVYNTRH---VCGFWGALITFSLSLFLYICSHVSTRALBCA 179
Db      152 SFPRYI--ALARMCSLFRFKHARLSCGLIW---MASVATLVPFAVHLQHNDACF 206
OY      180 KQONAAADATLVFIFYVVP--ALATLVAVLVLSRRRBDTLDLDYGLFESANR---L 234
Db      207 CFADVAEVOVWLEVTLCFIVPFAIIGLCYSLIVRLVRAH-----HRRG-LRPRKALCM 260
OY      235 LVATVCTQGLWTPHYLLLGHTVLIISRGK-----VDAYLGLHFKDPSKTLA 285
Db      261 ILAVLVFVFCMLPENVFISVH--LLQRTQPGACAPCKOSFRHAHL-TGHIV---NLAA 313
OY      286 FSSSFPTPLLYRYMNSPFSKQRLMK--LPCGDRHCSPDHMGVQVLT 332
Db      314 FNSNCINPLIYSFLGTFPRDKRLTYEQKTNLPALNRF--HAALKAVI 360

```

```

DT      15-JUN-1998 (Rel. 36, Created)
DT      15-JUN-1998 (Rel. 36, Last sequence update)
DT      05-JUN-2004 (Rel. 44, Last annotation update)
DE      Chemokine receptor-like 2 (G-protein coupled receptor GPR41).
GN      Name=cmkrl2; Synonyms=Gpr41;
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN=Sprague-Dawley; TISSUE=Lung;
RX      MEDLINE=97312546; PubMed=9168987;
RA      Bonini J.A., Anderson S.M., Steiner D.F.;
RT      Molecular cloning and tissue expression of a novel orphan G protein-
RT      coupled receptor from rat lung.
RL      Biochem. Biophys. Res. Commun. 234:190-193(1997).
CC      -1- FUNCTION: Orphan receptor; possibly for a chemokine.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -1- TISSUE SPECIFICITY: Highly expressed in brain and lung, lower
CC      expression in heart, skeletal muscle and kidney.
CC      -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; U92802; AAC53208.1; -.
DR      PIR; JC5509; JC5509.
DR      InterPro; IPR000276; GPCR_Rhodopsn.
DR      InterPro; IPR002286; P2_purinocptor.
DR      Pfam; PF00001; 7tm_1; 1.
DR      PRINTS; PR00237; GPCR_RHODOPSIN.
DR      PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR      PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW      G-protein coupled receptor; Glycoprotein; Transmembrane.
FT      DOMAIN 1 62
FT      TRANSMM 63 84
FT      DOMAIN 85 96
FT      TRANSMM 97 120
FT      DOMAIN 121 132
FT      TRANSMM 133 153
FT      DOMAIN 154 175
FT      TRANSMM 176 194
FT      DOMAIN 195 220
FT      TRANSMM 221 236
FT      DOMAIN 237 259
FT      TRANSMM 260 280
FT      DOMAIN 281 306
FT      TRANSMM 307 327
FT      DOMAIN 328 375
FT      DISULFID 130 207
FT      CARBOHYD 32 32
FT      CARBOHYD 44 44
SQ      SEQUENCE 375 AA; 42259 MW; 08B60576BDBE314 CRC64;

```

```

Query Match 13.5%; Score 233.5; DB 1; Length 375;
Best Local Similarity 24.6%; Pred. No. 1.7e-09;
Matches 85; Conservative 69; Mismatches 138; Indels 53; Gaps 16;

OY      12 LVEELPACDDLGSLSLGLGVGVPGCTYNALVLANLHASKASMTMPDVYVNNAV 71
Db      45 LMGDLSEHQGYVIAL-FLSCLYTIPLFPIGFVGNILIVVNISFREKKTIPDLFIINAA 103
OY      72 AGVLVLSALPVLHLPSSSRMALMSVGEVHALQIPFVNSLVAMYSTAL---LSLDH 127
Db      104 ADLILVADSLIEVFNHDEQYVIAVLCTFMSLFLQ-----NMYSVFELTWMSPDR 155
OY      128 YIERALPRTYMASVYNTRH---VCGFWGALITFSLSLFLYICSHVSTRALBCAKQN 183

```

```

Db 156 YL--ALAMRCRGLRRTKHAHLSGLIN---MASSATIVPPTAHLKRTTERACCPAD 210
Oy 184 ABAADATLVPIGYVP--ALATLVALVLSRVREDPTLDRGTGLEP---SAHRLVAT 238
Db 211 VREVOMLTEVLTGFIVPFALITGLCYSLIVRALIRAH-----RRHG-LRPRKALRMIFAV 264
Oy 239 VCTQGLMTPHVLLILGHVILSRCKPVDA-----HYGLL-HYVKPSKLLARSSSP 250
Db 265 VLVFETICWLPENVFTSVH--LLQWAPGGTTPCKQSFRAHYPLGLHIV---NLAFSNSNC 318
Oy 291 VTPLLRYMNOGSPPSKLAQRLMKR---LPGDRHCSPDHNGVOVL 332
Db 319 LSPLIYSFLGETFRDKLRUYAOKTSLPALNRF--HATLKAVI 360

RESULT 12
O9D392 PRELIMINARY; PRT; 375 AA.
AC O9D392.
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, last annotation update)
DE Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
DE enriched library, clones:6330420K13 product:CHEMOKINE RECEPTOR-LIKE 2
DE (G-PROTEIN COUPLED RECEPTOR GPR41) homolog.
GN Name=Gpr30;
OS Mus musculus (Mouse) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Mech. Enzymol. 303:19-44(1999).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 405:685-690(2001).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX SHIBATA K., Itoh M., Atzawa K., Nagaoka S., Saeki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Matsubiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;

```

RT	"RIKEN integrated sequence analysis (RISA) system-384-format
RT	sequencing pipeline with 384 multicapillary sequencer."
RN	Genome Res. 10:1757-1771(2000).
RP	[6]
RC	SEQUENCE FROM N.A.
RC	STRAIN-C57BL/6J; TISSUE=Medulla oblongata;
RA	Adechri J., Alzawa K., Akahira S., Akiyama T., Arai A., Aono H.,
RA	Arakawa T., Bono H., Carninci P., Fukuda S., Fukuishi Y., Furuno M.,
RA	Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiyao T., Hori P.,
RA	Iimotani K., Ishii Y., Itoh M., Iwano M., Kasukawa T., Kato H.,
RA	Kawai Y., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA	Matsuyama T., Miyazaki A., Nishii K., Nomura K., Numazaki R., Ohno M.,
RA	Otakeki Y., Okido T., Owa C., Salto K., Satoh H., Sakai C., Sakai K.,
RA	Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA	Sogabe Y., Suzuki H., Tagami M., Tegawa A., Takahashi F., Tanaka T.,
RA	Tajima Y., Toyota T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA	Muramatsu M., Hayashizaki Y.;
RL	Submitted (JUL-2000) to the EMBL/GenBank/DDbj databases.
CC	- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC	- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
DR	EMBL; AK018203; BABS118.1; -
DR	MGD; MG1:1924104; Gpr30.
DR	GO; GO:0016021; C: integral to membrane; IEA.
DR	GO; GO:0045028; F: putative nucleotide receptor activity, G- . . ; IEA.
DR	GO; GO:0004872; F: receptor activity; IEA.
DR	GO; GO:0001584; F: rhodopsin-like receptor activity; IEA.
DR	GO; GO:0007166; P: G-protein coupled receptor protein signaln. . ; IEA.
DR	InterPro; IPRO00276; GPCR_Rhodopsn.
DR	InterPro; IPRO02286; P2_purinoreceptor.
DR	Pfam; PF00001; 7tm_1_1_Purinoreceptor.
DR	PRINTS; PR00037; GPCRRHODOPSIN.
DR	PRINTS; PR01157; P2PURINOCEPTOR.
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR	PROSITE; PS02622; G_PROTEIN_RECEP_F1_2; 1.
KW	G-protein coupled receptor; Receptor; Transmembrane.
SO	SEQUENCE 375 AA; 42492 MW; AF7EA795C3AB695 CRC64;
Query Match	13.2%; Score 229.5; DB 2; Length 375;
Best Local Similarity	24.3%; Pred. No. 3.4e-09;
Matches 84; Conservative 70; Mismatches 138; Indels 53; Gaps 16;	
CY	12 LVEELPACQDLQLGLSLLSLGLVGVPGVCTVALLVNLNHSKASTMTMDPVFNNAV 71
DB	45 LTGOLSEHQYVIAL-FLSCLYTFLEPPIGFVGNGILLVNVISPREKKTIIDPLVINLAA 103
CY	72 AGVLISALAPRHLAGPPSSRWALMSVCGEYNVAIQIFNVSSLVAMYSTAL----LSLDH 127
DB	104 ADLLIVADSLIEVNDLBQYVDIADVLCFMSLFQI-----NMSSVFFLTMSBDR 155
CY	128 YIERBALPRTYMASVYNTKH---VCGFVWGALLTSPSSLYICSHYSTALCAKQN 183
DB	156 YL--ALKAKRGCGFRTHHARLSGILM--MASVSATLVPPFAVHLRHTEACCPCAD 210
CY	184 AEADADTLVFGIYVP--ALATLYAVLVLSVRREDTPLRDGTGRLE--SARLVAT 238
DB	211 VREVOVMLEVTGLFMFPAILGLCSLIRALIRAH-----KHKG-LPRRKOKALRMIPAV 264
CY	239 VCQDFGLMTPIRYLLLGHTVIISRGKPVDA-----HYGLL-HFYVDFEKLAFFSSF 290
DB	265 VLVPFICMLPENVAISH--LIQMTOEGDTPCKOSPRHAAPLTGHIV---NLAAFSNSC 318
CY	291 VTPLLRYNNOSPSPKLORLMKK--LPCCGRHSCPDHMGVOOYL 332
DB	319 LNPILVSPLEGTFRDKRLRVYEQTSPALNRFC---HATLKAVI 360
RESULT 13	
QBMP4	PRELIMINARY; PRT; 375 AA.
QBMP4	
QBMP4	
DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Mus musculus adult male pituitary gland cDNA, RIKEN full-length
 DE enriched library, clone:530404618 product:CHEMOKINE RECEPTOR-LIKE 2
 DE (G-PROTEIN COUPLED RECEPTOR GPR41) homolog.
 GN Name=Gpr30;
 OS Mus musculus (Mouse);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RP STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RC MEDLINE=99279253; PubMed=10396636;
 RX Carninci P.; Hayashizaki Y.;
 RA "High-efficiency full-length cDNA cloning";
 RL Mech. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RA the FANTOM Consortium,
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P.; Shibata Y.; Hayatsu N.; Sugahara Y.; Shibata K.; Itoh M.;
 RA Kono H.; Okazaki Y.; Muramatsu M.; Hayashizaki Y.;
 RT "Normalization and subraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K.; Itoh M.; Aizawa K.; Nagaoka S.; Sasaki N.; Carninci P.;
 RA Kono H.; Akiyama U.; Nishi K.; Kikunai T.; Tashiro H.; Itoh M.;
 RA Suni N.; Ishii Y.; Nakamura S.; Hazama M.; Nishine T.; Harada A.;
 RA Yanamoto R.; Matsumoto H.; Sakaguchi S.; Ikegami T.; Kaishogi K.;
 RA Fujiwaka S.; Inoue K.; Togawa Y.; Izawa M.; Ohara E.; Watabiki M.;
 RA Yoneda Y.; Ishikawa T.; Ozawa K.; Tanaka T.; Matsura S.; Kawai J.;
 RA Okazaki Y.; Muramatsu M.; Inoue Y.; Kita A.; Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RX Adachi U.; Aizawa K.; Akimura T.; Arakawa T.; Bono H.; Carninci P.;
 RA Fukuda S.; Furuno M.; Hanagaki T.; Hara A.; Hashizume W.;
 RA Hayashida K.; Hayatsu N.; Hiramoto K.; Hirooka T.; Hirozane T.;
 RA Hori F.; Imocani K.; Ishii Y.; Itoh M.; Kagawa I.; Kasukawa T.;
 RA Kawai H.; Kawai J.; Kojima Y.; Kondo S.; Kono H.; Kouda M.; Koya S.;
 RA Kurihara C.; Matsuyama T.; Miyazaki A.; Murata N.; Nakamura M.;
 RA Nishi K.; Nomura K.; Numazaki R.; Ohno M.; Ohsato N.; Okazaki Y.;
 RA Saito K.; Satoh H.; Sakai C.; Sakai K.; Sakazume N.; Sano H.;
 RA Sasaki D.; Shibata K.; Shinagawa A.; Shiraki T.; Sogabe Y.; Tagami M.;
 RA Tagawa A.; Takahashi F.; Takaku-Akahira S.; Takeda Y.; Tanaka T.;
 RA Tomaru A.; Toya T.; Yasunishi A.; Muramatsu M.; Hayashizaki Y.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 DR EMBL; AK030375; BAC26930.1; -.
 DR MGD; MGI:1924104; Gpr30.
 DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0045028; F:purinergic nucleotide receptor activity; G- . . ; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0001186; P:G-protein coupled receptor protein signalin. . . ; IEA.
 DR InterPro: IPR002376; GPCR_Rhodopsn.
 DR InterPro: IPR002286; P2_purinocptr.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS; PR00227; GPCR_RHODOPSIN.
 DR PRINTS; PR01157; P2Y_PURINOCPTR.
 DR PROSITE; PS00237; G_PROTEIN_RECPT_P1_1; 1.
 DR PROSITE; PS00262; G_PROTEIN_RECPT_P1_2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 375 AA; 42444 MW; A5E7D795C3A6AB05 CRC64;
 Query Match 12.9%; Score 223.5; DB 2; Length 375;
 Best Local Similarity 24.1%; Pred. No. 3,7e-09;
 Matches 83; Conservative 70; Mismatches 139; Indels 53; Gaps 16;
 QY 12 LVEELPACODQLGLSLILGLVGVPGVGLCYNALVLVLAHLSKASMTDPVFNNAV 71
 DB 45 LFGDLSHQVYIAL-FLSCLYTIPLFPIGFGNILLIVNISPEKKTIPDLYFINIAA 103
 QY 72 AGVLVSLAPVHLDPSSRMALSVGSEHVVALDQIPFNVSLSVMYSTAL----LSLDH 127
 DB 104 ADLLIVADSLIEVFNLDEQYDIAVLCTFMSLFLQI-----NMYSVFPLTWMSFDR 155
 QY 128 YIERLALPRYMASVYNTSH---VCGFWGAGLTPSSILFYICSHVSTRALBCAKQN 183
 DB 156 YL--ALAKAMRCGLFRTKHARLSCGLIW--MASVSATLVFTVAHLRHTEAFCRAD 210
 QY 184 AEAADATLVFIQVVP--ALATLVALLSRVRBDTPLDRTGSLER---SAHRLVAT 238
 DB 211 VAEVQMLEVTLDGFIMPFAIIGCYSLIVALLRAH-----RRRG-LRPRRQALMIRFV 264
 QY 239 VCTQGLMTPHYLLILGTVITTSRQKPDV-----HYGLL-HFVDFSLAFSSSF 290
 DB 265 VLVFFICMLPEVNTFSVH--LLQWTPQGPCKQSFRIHAYPLTGHIV---MLAITSNC 318
 QY 291 VVPLLYRVNNGSPFSLQRLMKK--LPCGDHSCPDHMGVOQL 332
 DB 319 LNEPLYSFLGETFRDRLVVEQKTSLPALNFC--HATLKAVI 360
 RESULT 14
 OBHZN5 PRELIMINARY; PRT; 355 AA.
 ID QGHZNS;
 AC QGHZNS;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Interleukin 8 receptor B CXCR2.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 ON NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA HORTICK R.A.; Zhao J.; Swanson R.N.; Webb M.L.; Strohl B.;
 RA Baldwin U.J.; Auld D.S.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 DR EMBL; AF540792; AA017318.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004918; F:interleukin-8 receptor activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0006335; P:chemotaxis; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . . ; IEA.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR000057; IL8R_Rhodopsn.
 DR InterPro: IPR000174; IL8R_receptor.

DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPN.
 DR PRINTS; PR00427; INTRLEUKIN8R.
 DR PRINTS; PR00573; INTRLEUKIN8R.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; 1.
 DR PROSITE; PS00262; G_PROTEIN_RECEPTOR_F2_1; 1.
 DR G-protein coupled receptor; Receptor; Transmembrane.
 KM GSEQUENCE 355 AA; 40292 MW; 6A9BP9A921F470C5 CRC64;

Query Match 12.0%; Score 208.5; DB 2; Length 355;
 Best Local Similarity 23.5%; Pred. No. 1.2e-07;
 Matches 84; Conservative 59; Mismatches 127; Indels 87; Gaps 13;

QY 1 MNSCSWNGTGVLELP-----ACQDLQGLS-----LLSLGLVGVGVG 41
 DB 9 LMGEDPSNYSDDLPPSLPDVAPCRPESLEINKYFVVIITVALVPLSLGL----- 60
 QY 42 LCNALVLNLHSHKASMTMPDVFVNMAVAGLVLSALAPVHLGPPSRWALMSVGGEV 101
 DB 61 --NSLVMVLVHSHRVGRSITDYLNLNADLLPALTLPI-----MAAKVGMW 108
 QY 102 H-----VALQIPFNVSLSVAMYSTALLSDHYIERA-LPRTYMASVYNTRHVCGFVWG 153
 DB 109 FGTELCRVSLKENVFYSGILL--LACISVDRYLAIVHATRTLQKRYLVKFCLSIWS 166
 QY 154 GALLTPSSLLFYICSHVSTRALECAGKQNAEADATLVF-----IGYVPALATYAL 207
 DB 167 LSLLLAPVLLPRRAVYPPYISPVCEYDMGNTAKMWMVRIILPQTGFIVPLIMLPCY 226
 QY 208 VLSRVREDTPLDRTGRLEPSAHRLLVAVTCTQPGMLTPHYLLIGHTVIISR----- 262
 DB 227 GLTLR-----TLRAHNGQ-KHRAKRVIPAVVLIFLLCMLPYHLVLAADTLMTRTLINET 280
 QY 263 -----GKPVDA-HYLGILLFVNDPSKLLAFSSSFVTPLLYRYNQSPPSKLQRLM 311
 DB 281 CORRNIDQALDATEILGILH-----SCNPLIYAPIGQKFRHGLLKITL 324

RESULT 15

Q8HZN3 PRELIMINARY; PRT; 355 AA.

AC Q8HZN3;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
 DE Interleukin 8 receptor B CXCR2.
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 NC NCBI_TaxID=9557;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Horlick R.A., Zhao J., Swanson R.N., Webb M.L., Strohl B.,
 RA Baldwin J.J., Auld D.S.,
 CC Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
 CC - SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC - SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 DR EMBL; AF540794; AAI17320.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004918; F:interleukin-8 receptor activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0006935; P:chemotaxis; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopn.
 DR InterPro; IPR000057; IL8_receptor.
 DR InterPro; IPR000174; IL8_receptor.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPN.
 DR PRINTS; PR00427; INTRLEUKIN8R.
 DR PRINTS; PR00573; INTRLEUKIN8R.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; 1.

DR PROSITE; PS00262; G_PROTEIN_RECEPTOR_F2_1; 1.
 KM G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 355 AA; 40253 MW; 69BDF0C0D0717A53 CRC64;

Query Match 11.9%; Score 206.5; DB 2; Length 355;
 Best Local Similarity 23.5%; Pred. No. 1.8e-07;
 Matches 81; Conservative 61; Mismatches 138; Indels 65; Gaps 12;

QY 2 WSCSWFNGTGVLELP-----ACQDLQGLS-----LLSLGLVGVGVG 42
 DB 10 WTGEDPSNYSDDLPPSLPDVAPCRPESLEINKYFVVIITVALVPLSLGL----- 60
 QY 43 CYNALVLNLHSHKASMTMPDVFVNMAVAGLVLSALAPVHLGPPSRWALMSVGGEV 102
 DB 61 --NSLVMVLVHSHRVGRSITDYLNLNADLLPALTLPI-----MAAKVGMW 109
 QY 103 -----VALQIPFNVSLSVAMYSTALLSDHYIERA-LPRTYMASVYNTRHVCGFVWG 154
 DB 110 GFTELCRVSLKENVFYSGILL--LACISVDRYLAIVHATRTLQKRYLVKFCLSIWL 167
 QY 155 ALLTPSSLLFYICSHVSTRALECAGKQNAEADATLVF-----IGYVPALATYAL 208
 DB 168 LSLLLAPVLLPRRAVYPPYISPVCEYDMGNTAKMWMVRIILPQTGFIVPLIMLPCY 227
 QY 209 VLSRVREDTPLDRTGRLEPSAHRLLVAVTCTQPGMLTPHYLLIGHTVIISR--GKPV 266
 DB 228 GLTLR-----TLRAHNGQ-KHRAKRVIPAVVLIFLLCMLPYHLVLAADTLMTRTLINET 281
 QY 267 DAHYLGILLFVNDPSKLLAFSSSFVTPLLYRYNQSPPSKLQRLM 311
 DB 282 QRR--SDINQALDATEILGIFHSCNPLIYAPIGQKFRHGLLKITL 324

Search completed: December 14, 2004, 20:25:06
 Job time : 57.3673 secs

This Page Blank (uspto)

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. AUTHORS 1 Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R., Elliott, V.S., Ramkumar, J., Baughn, M.R., Kallick, D.A., Walla, N.K., Haila, A.U., Yao, M.G., Lu, Y., Tribouley, C.M., Policky, J.L., Kearney, L., Graul, R.C., Warren, B.A. and Ding, L. G-protein coupled receptors Patent: WO 0210387-A 25 07-FEB-2002; Incyte Genomics, Inc. (US) FEATURES Location/Qualifiers 1. .1130 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="Incyte ID No: 7474767CB1" ORIGIN Query Match 100.0%; Score 27; DB 6; Length 1130; Best Local Similarity 100.0%; Pred. No. 0.5; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 1 GCGCTCATGAGACACGACGCCAC 27 6 GCGCTCATGAGACACGACGCCAC 32 RESULT 3 AY635179 1167 bp mRNA linear PRI 04-JUL-2004 LOCUS Homo sapiens G protein-coupled receptor 139 (GPR139) mRNA, complete cds. VERSION AY635179 GI:49413551 SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1167) Bonner, T.I., Nagle, J.W. and Kaufman, D. Complete coding sequence of GPR139 Unpublished 2 (bases 1 to 1167) Bonner, T.I., Nagle, J.W. and Kaufman, D. Direct Submission Submitted (25-MAY-2004) Lab of Genetics, NIMH, Bldg 36, Rm 3D06, MSC4094, Bethesda, MD 20892-4094, USA FEATURES Location/Qualifiers 1. .1167 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="16" /map="16p12.3" /issue_type="brain" 1. .1167 /gene="GPR139" 92. .1153 /gene="GPR139" /codon_start=1 /product="G protein-coupled receptor 139" /protein_id="AA165818.1" /db_xref="GI:49413552" /translation="MEHTRAHLAANSLSWSPGACGLGVVVVYYSILLCGLPAN ILTVILSQLVARQKSSVNYLLAADIIVLFLVVDLLEDFILNMQPOVDEK IIEVLFSSHTSIWITVPLTIDRYANCHPKHTVSPARTKRVTSVITCGLTS IPYVWPNIWTEVDYISTVHVLLVWCHFTVYVPCSLFFILNSTVYKLRKSNFRL GYSTGKTAIFITTSIFALIMAPRIIMIIYHVGAPIONRWLVHMSDINMAL NTAINFLYCFISKRFRMAATLKAFFCKQKOPQVFTNNHFSITSSPWISPSNSHC IKMLVYQDKNGKPIKVP" complement (1142..1167) primer_bind ORIGIN

Query Match 100.0%; Score 27; DB 9; Length 1167; Best Local Similarity 100.0%; Pred. No. 0.5; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 1 GCGCTCATGAGACACGACGCCAC 27 86 GCGCTCATGAGACACGACGCCAC 112 Db 6 GCGCTCATGAGACACGACGCCAC 112 RESULT 4 AX664701 1526 bp DNA linear PAT 22-MAR-2003 LOCUS Sequence 10 from Patent WO02074960. AX664701 AX664701 VERSION AX664701.1 GI:29164461 KEYWORDS SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1526) Leiby, K.R., Kapeller-Libermann, R. and Gluckmann, M. 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions of human proteins and uses thereof Patent: WO 02074960-A 10 26-SEP-2002; Millennium Pharmaceuticals, Inc. (US) FEATURES Location/Qualifiers 1. .1526 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 139. .1200 /note="unnamed protein product" /codon_start=1 /protein_id="CAD80046.1" /db_xref="GI:29164462" /translation="MEHTRAHLAANSLSWSPGACGLGVVVVYYSILLCGLPAN ILTVILSQLVARQKSSVNYLLAADIIVLFLVVDLLEDFILNMQPOVDEK IIEVLFSSHTSIWITVPLTIDRYANCHPKHTVSPARTKRVTSVITCGLTS IPYVWPNIWTEVDYISTVHVLLVWCHFTVYVPCSLFFILNSTVYKLRKSNFRL GYSTGKTAIFITTSIFALIMAPRIIMIIYHVGAPIONRWLVHMSDINMAL NTAINFLYCFISKRFRMAATLKAFFCKQKOPQVFTNNHFSITSSPWISPSNSHC IKMLVYQDKNGKPIKVP" ORIGIN Query Match 100.0%; Score 27; DB 6; Length 1526; Best Local Similarity 100.0%; Pred. No. 0.5; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 1 GCGCTCATGAGACACGACGCCAC 27 Db 133 GCGCTCATGAGACACGACGCCAC 159 RESULT 5 AC021089/c 101882 bp DNA linear PRI 22-MAR-2003 LOCUS Homo sapiens chromosome 16 clone CTD-2264D9, complete sequence. AC021089 AC021089 VERSION AC021089.8 GI:29150340 KEYWORDS HTG. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 101882) DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory. Direct Submission Unpublished 2 (bases 1 to 101882) DOE Joint Genome Institute. ORIGIN

TITLE Direct Submission
JOURNAL Submitted (14-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 5 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Mar 22, 2003 this sequence version replaced gi:20330797.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.sbgsc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
Location/Qualifiers
1. 101882
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTD-2264D9"

ORIGIN
Query Match 100.0%; Score 27; DB 9; Length 101882;
Best Local Similarity 100.0%; Pred. No. 0.45; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCCGAC 27
|||||
Db 60121 GCGCTCATGAGCAGCAGCCGAC 60095

RESULT 6
AL139235 0/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AL139235 Accession AL139235
Fragment Name Begin End
AL139235_0 1 110000
AL139235_1 100001 210000
AL139235_2 200001 310000
AL139235_3 300001 407025
LOCUS AL139235 407025 bp DNA 11near HTG 16-JUN-2001
DEFINITION Homo sapiens chromosome 9 clone RP11-3215, 31 unordered pieces.
ACCESSION AL139235
VERSION AL139235.22 GI:11034497
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE 1 Plumb, B.
AUTHORS Direct Submission
TITLE Submitted (14-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 15A, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Oct 26, 2000 this sequence version replaced gi:11024977.
----- Genome Center
Center: Sanger Centre
Center code: SC

Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: B3215
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 389894 bases at least Q40
Consensus quality: 395933 bases at least Q30
Consensus quality: 400050 bases at least Q20
Insert size: 14695; 17.7% error; agarose-fp
Quality coverage: 4.93x in Q20 bases; sum-of-confids Quality
coverage: 14.48x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2426 2425: contig of 2425 bp in length
* 2426 2525: gap of 100 bp
* 2526 10023: contig of 7498 bp in length
* 10024 10123: gap of 100 bp
* 10124 36364: contig of 26241 bp in length
* 36365 36464: gap of 100 bp
* 36465 43873: contig of 7409 bp in length
* 43874 43973: gap of 100 bp
* 43974 49988: contig of 6015 bp in length
* 49989 50088: gap of 100 bp
* 50089 93603: contig of 43515 bp in length
* 93604 93703: gap of 100 bp
* 93704 96191: contig of 2488 bp in length
* 96192 96291: gap of 100 bp
* 96292 99433: contig of 3142 bp in length
* 99434 99533: gap of 100 bp
* 99534 101556: contig of 2023 bp in length
* 101557 101656: gap of 100 bp
* 101657 184109: contig of 82453 bp in length
* 184110 184209: gap of 100 bp
* 184210 290696: contig of 106487 bp in length
* 290697 290796: gap of 100 bp
* 290797 292892: contig of 2096 bp in length
* 292893 292992: gap of 100 bp
* 292993 296018: contig of 3026 bp in length
* 296019 296118: gap of 100 bp
* 296119 322154: contig of 26036 bp in length
* 322155 322254: gap of 100 bp
* 322255 324646: contig of 2392 bp in length
* 324647 324746: gap of 100 bp
* 324747 329059: contig of 4323 bp in length
* 329060 329159: gap of 100 bp
* 329160 331371: contig of 2202 bp in length
* 331372 331471: gap of 100 bp
* 331472 333909: contig of 2438 bp in length
* 333910 334009: gap of 100 bp
* 334010 369445: contig of 35436 bp in length
* 369446 369545: gap of 100 bp
* 369546 371996: contig of 2451 bp in length
* 371997 372096: gap of 100 bp
* 372097 375925: contig of 3829 bp in length
* 375926 376025: gap of 100 bp
* 376026 378132: contig of 2107 bp in length
* 378133 378232: gap of 100 bp
* 378233 381590: contig of 3338 bp in length
* 381591 381690: gap of 100 bp
* 381691 383726: contig of 2036 bp in length
* 383727 383826: gap of 100 bp
* 383827 390832: contig of 7006 bp in length

```
* 39083 39032: gap of 100 bp
* 39033 39416: contig of 3284 bp in length
* 39417 39416: gap of 100 bp
* 394317 396358: contig of 2042 bp in length
* 396359 396458: gap of 100 bp
* 396459 398733: contig of 2275 bp in length
* 398734 398833: gap of 100 bp
* 398834 402385: contig of 3552 bp in length
* 402386 402485: gap of 100 bp
* 402486 404522: contig of 2037 bp in length
* 404523 404623: gap of 100 bp
* 404623 407025: contig of 2403 bp in length.
FEATURES
    source
        1..407025
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="9"
            /clone="RPL1-3215"
            /clone_id="RPC1-11.1"
            1..2425
                /note="assembly_fragment:05385
                fragment_chain:1"
                2528..10023
                    /note="assembly_fragment:03131
                    fragment_chain:1"
                    10124..36364
                        /note="assembly_fragment:06402
                        fragment_chain:1"
                        36465..43873
                            /note="assembly_fragment:01436
                            fragment_chain:1"
                            43974..49988
                                /note="assembly_fragment:04547
                                fragment_chain:1"
                                50089..93603
                                    /note="assembly_fragment:01636
                                    fragment_chain:1"
                                    93704..96191
                                        /note="assembly_fragment:02957
                                        fragment_chain:2"
                                        96292..99433
                                            /note="assembly_fragment:00499
                                            fragment_chain:2"
                                            99534..101556
                                                /note="assembly_fragment:03112
                                                fragment_chain:2"
                                                101657..184109
                                                    /note="assembly_fragment:01419
                                                    fragment_chain:5"
                                                    184210..290696
                                                        /note="assembly_fragment:00925
                                                        fragment_chain:3"
                                                        290797..292892
                                                            /note="assembly_fragment:03065
                                                            fragment_chain:4"
                                                            292993..296018
                                                                /note="assembly_fragment:05057
                                                                fragment_chain:4"
                                                                296119..322154
                                                                    /note="assembly_fragment:00269"
                                                                    fragment_chain:4"
                                                                    322255..324646
                                                                        /note="assembly_fragment:00866"
                                                                        fragment_chain:4"
                                                                        324747..329069
                                                                            /note="assembly_fragment:01113"
                                                                            fragment_chain:4"
                                                                            329170..331371
                                                                                /note="assembly_fragment:01461"
                                                                                fragment_chain:4"
                                                                                331472..333909
                                                                                    /note="assembly_fragment:01516"
                                                                                    fragment_chain:4"
                                                                                    334010..369445
                                                                                        /note="assembly_fragment:01942.0"
                                                                                        fragment_chain:4"
                                                                                        369346..371996
                                                                                            /note="assembly_fragment:02194"
```

```
misc_feature 372097..375925
              /note="assembly_fragment:02249"
              376026..378132
                  /note="assembly_fragment:02485"
                  378233..381590
                      /note="assembly_fragment:02649"
                      381691..383726
                          /note="assembly_fragment:02869"
                          383827..390833
                              /note="assembly_fragment:03540"
                              390933..394216
                                  /note="assembly_fragment:04405"
                                  394317..396358
                                      /note="assembly_fragment:04595"
                                      396459..398733
                                          /note="assembly_fragment:05168"
                                          398834..402385
                                              /note="assembly_fragment:06106"
                                              402486..404522
                                                  /note="assembly_fragment:06204"
                                                  404623..407025
                                                      /note="assembly_fragment:06292"
```

ORIGIN

```
Query Match 100.0%; Score 27; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GGGCTCATGAGACACGACGCCAC 27
Db 76457 GGGCTCATGAGACACGACGCCAC 76431
```

```
RESULT 7
AC119798 140885 bp DNA linear HTG 06-JAN-2004
LOCUS Rattus norvegicus chromosome 1 clone RP32-329N19 map q35, ***
DEFINITION AC119798.2 GI:40646047
ACCESSION AC119798.2
VERSION AC119798.2
KEYWORDS HTG; HTGS PHASE2; HTGS FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 140885)
Taezenzer,S., Monti,J., Gloeckner,G., Goesele,C., Baumgart,C.,
Huebner,N. and Platzzer,M.
Rat chromosome 1 genomic sequence
Unpublished
2 (bases 1 to 140885)
Taezenzer,S. and Platzzer,M.
Direct Submission
Submitted (02-MAY-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
3 (bases 1 to 140885)
Lagemann,D. and Platzzer,M.
Direct Submission
Submitted (06-JAN-2004) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
On Jan 6, 2004 this sequence version replaced gi:20389312.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: RA23
Center clone name: Rp32-329N19
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

Assembly program: Phrap; version 0.990329
 Consensus quality: 140564 bases at least Q40
 Consensus quality: 140589 bases at least Q30
 Consensus quality: 140641 bases at least Q20
 Quality coverage: 12.80x

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

- * NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
- * This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
- * 1 27033: contig of 27033 bp in length
- * 27034 27133: gap of unknown length
- * 27134 67824: contig of 40691 bp in length
- * 67825 67924: gap of unknown length
- * 67925 140885: contig of 72961 bp in length.

FEATURES

source

1.140885
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /chromosome="1"
 /map="q35"
 /clone="RP32-329N19"
 /clone_1b="RP32-32 BAC library"

ORIGIN

Query Match 100.0%; Score 27; DB 2; Length 140885;
 Best Local Similarity 100.0%; Pred. No. 0.45;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCCGAC 27

DB 117161 GCGCTCATGAGCAGCAGCCGAC 117135

RESULT 8
 AC124438 183326 bp DNA linear ROD 13-NOV-2003

LOCUS
 AC124438 Mus musculus BAC clone RP24-260K9 from chromosome 7, complete

DEFINITION

AC124438 183326 bp DNA linear ROD 13-NOV-2003

ACCESSION

AC124438 GI:23462943

VERSION

AC124438.4 GI:23462943

KEYWORDS

HTG.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

Bukacinski, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

McLellan, M., Bielicki, L., Dignau, G. and Doeber, A.

TITLE

The sequence of Mus musculus BAC clone RP24-260K9

JOURNAL

Unpublished (2001)

REFERENCE

2 (bases 1 to 183326)

AUTHORS

Wilson, R.

TITLE

Sequencing of Mus musculus

JOURNAL

Unpublished (2001)

REFERENCE

3 (bases 1 to 183326)

AUTHORS

McPherson, J.D. and Waterston, R.H.

TITLE

Direct Submission

JOURNAL

Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park

REFERENCE

4 (bases 1 to 183326)

AUTHORS

McPherson, J.D. and Waterston, R.H.

TITLE

Direct Submission

JOURNAL

Submitted (01-SEP-2002) Genome Sequencing Center, 4444 Forest Park

REFERENCE

5 (bases 1 to 183326)

AUTHORS

McPherson, J.D. and Waterston, R.H.

TITLE

Direct Submission

JOURNAL

Submitted (02-OCT-2002) Genome Sequencing Center, 4444 Forest Park

REFERENCE

6 (bases 1 to 183326)

AUTHORS

Wilson, R.

TITLE

Direct Submission

JOURNAL

Submitted (13-NOV-2003) Department of Genetics, Washington

REFERENCE

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

AUTHORS

On Oct 2, 2002 this sequence version replaced gi:22539265.

TITLE

Genome Center

COMMENT

Center: Washington University Genome Sequencing Center
 Center code: WUSGC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@wustl.wustl.edu
 Center project name: M_BB0260K09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wee Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RP32-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source

1.183326
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="7"
 /map="7"
 /clone="RP24-260K9"
 /clone_1b="RP32-24"
 421..527
 /rpt_family="L1"
 780..849
 /rpt_family="L1"
 1229..1489
 /rpt_family="L1"
 1925..2147
 /rpt_family="MER1_type"
 2269..2345
 /rpt_family="B4"

```
repeat_region      2432. .2639      /rpt_family="L1"
repeat_region      2657. .2747      /rpt_family="L1"
repeat_region      2753. .2819      /rpt_family="ID"
repeat_region      2825. .2891      /rpt_family="L1"
repeat_region      3225. .4557      /rpt_family="L1"
repeat_region      5567. .5560      /rpt_family="MIR"
repeat_region      6490. .6682      /rpt_family="B2"
repeat_region      6777. .6881      /rpt_family="Alu"
repeat_region      6889. .7045      /rpt_family="B4"
repeat_region      9156. .9335      /rpt_family="MER2_type"
repeat_region      10113. .10156      /rpt_family="B4"
repeat_region      10440. .10701      /rpt_family="B4"
repeat_region      11049. .11114      /rpt_family="ID"
repeat_region      11205. .11581      /rpt_family="L1"
repeat_region      12639. .12861      /rpt_family="B4"
repeat_region      12916. .12988      /rpt_family="B4"
repeat_region      13314. .13598      /rpt_family="B4"
repeat_region      14985. .15253      /rpt_family="B4"
repeat_region      15284. .15355      /rpt_family="MIR"
repeat_region      15669. .15726      /rpt_family="ERV1"
repeat_region      16139. .16229      /rpt_family="L1"
repeat_region      16418. .16525      /rpt_family="L1"
repeat_region      16880. .16932      /rpt_family="L1"
repeat_region      16942. .17139      /rpt_family="MIR"
repeat_region      17302. .17445      /rpt_family="L1"
repeat_region      17458. .17508      /rpt_family="Alu"
repeat_region      17667. .17787      /rpt_family="L1"
repeat_region      18127. .18260      /rpt_family="Alu"
repeat_region      18386. .18472      /rpt_family="L2"
repeat_region      18473. .18818      /rpt_family="MIR"
repeat_region      18831. .19229      /rpt_family="MALR"
repeat_region      19435. .19530      /rpt_family="MALR"
repeat_region      20110. .20262      /rpt_family="MIR"
repeat_region      20395. .20447      /rpt_family="B4"
repeat_region      20458. .20581      /rpt_family="ERV1"
repeat_region      21632. .21854      /rpt_family="MALR"
repeat_region      21927. .22288      /rpt_family="MER1_type"
repeat_region
```

```
repeat_region      /rpt_family="MALR"
repeat_region      23000. .23102      /rpt_family="MIR"
repeat_region      23289. .23480      /rpt_family="B2"
repeat_region      23481. .23670      /rpt_family="B2"
repeat_region      23989. .24034      /rpt_family="B2"
repeat_region      24461. .24730      /rpt_family="L2"
repeat_region      25760. .25970      /rpt_family="B4"
repeat_region      26406. .26524      /rpt_family="MER1_type"
repeat_region      26748. .27090      /rpt_family="Alu"
repeat_region      28283. .28734      /rpt_family="MALR"
repeat_region      29104. .29319      /rpt_family="ERV1"
repeat_region      29316. .29425      /rpt_family="MER2_type"
repeat_region      29597. .29747      /rpt_family="MER2_type"
repeat_region      29825. .29939      /rpt_family="MIR"
repeat_region      30206. .30399      /rpt_family="MIR"
repeat_region      31148. .31285      /rpt_family="MER1_type"
repeat_region      32032. .32336      /rpt_family="MER1_type"
repeat_region      /rpt_family="MALR"
```

```
Query Match      100.0%; Score 27; DB 10; Length 183326;
Best Local Similarity 100.0%; Pred. No. 0.44; Mismatches 0; Indels 0; Gaps 0;
Matches 27; Conservative
```

```
QY      1 GCGCTCATGAGACACGACGCCCCAC 27
Db      66628 GCGCTCATGAGACACGACGCCCCAC 66654
```

```
RESULT 9
AC147502      202934 bp      DNA      linear      ROD 15-MAY-2004
LOCUS      Mus musculus BAC clone RP23-111N9 from chromosome 7, complete
DEFINITION
ACCESSION      AC147502
VERSION      AC147502.2 GI:46275996
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 202934)
Swearengen-Shahid,S., Shahid,S., Cotton,M., Kozlowski,A. and
Meyer,R.
The sequence of Mus musculus BAC clone RP23-111N9
Unpublished (2001)
2 (bases 1 to 202934)
Wilson,R.K.
Direct Submission
Submitted (05-DEC-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 202934)
Wilson,R.K.
Direct Submission
Submitted (08-APR-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 202934)
Wilson,R.K.
Direct Submission
```


JOURNAL

Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
On Apr 8, 2004 this sequence version replaced gi:38708154.

COMMENT

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.wustl.edu
----- Summary Statistics
Center project name: M_BA0111N09

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence, and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Meg Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC library has been constructed by Kazuhiro Osegawa and Minako Tateno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC124438.

FEATURES

source
1. .202934
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/map="7"
/clone="RP23-111N9"
/clone_1tb="RPCI-23"
1959. .2627
repeat_region
/rpt_family="L1"
2691. .3205
repeat_region
/rpt_family="L1"
3505. .3950
repeat_region
/rpt_family="L1"
4204. .4359
repeat_region
/rpt_family="L1"
5371. .5662
repeat_region
/rpt_family="B4"
5767. .6129
repeat_region
/rpt_family="ERV1"
6738. .6796
repeat_region
/rpt_family="ERV1"
6770. .6830
repeat_region
/rpt_family="ERV1"
6874. .6946
repeat_region
/rpt_family="ERV1"
6990. .7039
repeat_region
/rpt_family="ERV1"
7040. .7314
repeat_region
/rpt_family="B4"
7901. .7986
repeat_region

repeat_region
/rpt_family="ERV1"
8120. .8340
/rpt_family="MALR"
8363. .8545
unsure
/note="Sequence derived from one plasmid subclone."
repeat_region
11282. .11670
/rpt_family="MALR"
12408. .12698
repeat_region
/rpt_family="B4"
13934. .14144
repeat_region
/rpt_family="B4"
14154. .14294
repeat_region
/rpt_family="ALU"
15089. .15228
repeat_region
/rpt_family="MALR"
15259. .15531
repeat_region
/rpt_family="B4"
15907. .16027
repeat_region
/rpt_family="B4"
16070. .16180
repeat_region
/rpt_family="MIR"
17754. .17867
repeat_region
/rpt_family="ALU"
19213. .19396
repeat_region
/rpt_family="MER1_type"
19774. .20083
repeat_region
/rpt_family="MALR"
20358. .20449
repeat_region
/rpt_family="L2"
20624. .20721
repeat_region
/rpt_family="MER2_type"
20791. .20941
repeat_region
/rpt_family="B2"
21006. .21133
repeat_region
/rpt_family="B4"
22495. .22732
repeat_region
/rpt_family="B4"
22897. .22990
repeat_region
/rpt_family="MER1_type"
23022. .23100
repeat_region
/rpt_family="ID"
23208. .23498
repeat_region
/rpt_family="B4"
23850. .23934
repeat_region
/rpt_family="MIR"
25280. .25406
repeat_region
/rpt_family="B4"
25682. .25852
unsure
/note="Unresolved simple sequence repeat."
repeat_region
26200. .26388
/rpt_family="MER1_type"
26659. .26898
repeat_region
/rpt_family="B4"
28500. .28665
repeat_region
/rpt_family="B4"
28668. .28810
repeat_region
/rpt_family="TC2"
29284. .29461
repeat_region
/rpt_family="B4"
29469. .29542
repeat_region
/rpt_family="MER1_type"
31411. .31526
repeat_region
/rpt_family="ERV1"
33813. .33868
repeat_region
/rpt_family="B4"
33897. .34003
repeat_region
/rpt_family="ALU"
35786. .35962
repeat_region
/rpt_family="L1"
35988. .36388
repeat_region
/rpt_family="MALR"
36433. .36586
repeat_region
/rpt_family="L1"

```

repeat_region      36698..36835
                    /rpt_family="MALR"
repeat_region      37663..37957
                    /rpt_family="L1"
repeat_region      37971..38090
                    /rpt_family="Alu"
repeat_region      38143..39094
                    /rpt_family="L1"
repeat_region      39871..40127
                    /rpt_family="RNER17C"
repeat_region      40375..40589
                    /rpt_family="B2"
repeat_region      40617..40715
                    /rpt_family="ID"
repeat_region      40978..41128
                    /rpt_family="MR"
repeat_region      43035..43298
                    /rpt_family="L1"
repeat_region      43928..44131
                    /rpt_family="L2"
repeat_region      44422..44452
                    /rpt_family="L2"
repeat_region      44463..44862
                    /rpt_family="RNER17C"
repeat_region      44655..44914
                    /rpt_family="ERRV"

```

```

Query Match      100.0%; Score 27; DB 10; Length 202934;
Best Local Similarity 100.0%; Pred. NO. 0.44;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 GCGCTCATGGAGCAGCAGCGCCAC 27
Db      190319 GCGCTCATGGAGCAGCAGCGCCAC 190345

```

```

RESULT 10
AC114097/c      230813 bp      DNA      linear      HTG 13-MAY-2003
LOCUS
AC114097
AC114097.5 GI:30580092
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 230813)
Muzny,D,Matzke,M, Metzker,M, Lee,A, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Cencier,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,U,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,J, D'Souza,L,
Davila,M,L, Davis,S, Davy-Carroll,L, De Anda,C, Dederich,D,
Delgado,O, Denson,S, Derramo,C, Ding,Y, Dinh,H, Divya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Eugene,C, Evans,C, A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Frazer,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Gebreyes,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guayana,W,
Guerache,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,
Hollins,B, Howells,S, Hulik,S, Hume,J, Idubidiz,D, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovac,C,
Kowalski,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,

```

REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:24819108.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLTA
Center clone name: CH230-90E1
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 226424 bases at least Q40
Consensus quality: 227854 bases at least Q30
Consensus quality: 228799 bases at least Q20
Estimated insert size: 240103; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_difac.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 230813: contig of 230813 bp in length.
Location/Qualifiers

FEATURES

Source

1. 230813
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-90E1"

ORIGIN

Query Match 100.0%; Score 27; DB 2; Length 230813;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCGCTCATGGAGCAGCAGCGCCAC 27
Db 90822 GCGCTCATGGAGCAGCAGCGCCAC 90796

RESULT 11

AP006494.2/c
WPCOMMENT
Sequence split into 9 fragments LOCUS AP006494 Accession AP006494

Fragment Name	Begin	End
AP006494.0	1	110000
AP006494.1	100001	210000
AP006494.2	200001	310000
AP006494.3	300001	410000
AP006494.4	400001	510000
AP006494.5	500001	610000
AP006494.6	600001	710000
AP006494.7	700001	810000
AP006494.8	800001	849694

Continuation (3 of 9) of AP006494 from base 200001 (AP006494 Cyanidioschyzon merolae str.

Query Match 84.4%; Score 22.8; DB 2; Length 110000;
Best Local Similarity 92.3%; Pred. No. 29;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GCGCTCATGGAGCAGCAGCGCCCA 26
Db 64886 GCGCTCATGGAGCAGCAGCGCCCA 64861

RESULT 12

BD183201 32 bp DNA linear PAT 17-JUN-2003

LOCUS BD183201 Novel G protein-coupled receptor and its DNA.

DEFINITION BD183201

ACCESSION BD183201.1 GI:31875401

VERSION JP 2002345481-A/1.

KEYWORDS JP 2002345481-A/1.

SOURCE Synthetic construct

ORGANISM Artificial sequences.

REFERENCE 1 (bases 1 to 32)

AUTHORS Terao, Y., Matsui, H. and Shintani, Y.

TITLE Novel G protein-coupled receptor and its DNA

JOURNAL Patent: JP 2002345481-A 1 03-DEC-2002;

COMMENT TAKEDA CHEMICAL INDUSTRIES LTD

OS Artificial Sequence

PN JP 2002345481-A/1

PD 03-DEC-2002

PR 01-JUN-2001 JP 2001166688

PI YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI

PC C12N15/09, A61K45/00, A61P1/00, A61P9/00, A61P25/00, PC

A61P29/00, PC

A61P35/00, A61P37/00, C07K14/705, C07K16/28, C12N1/15, C12N1/19, PC

C12N1/21,

PC C12N5/10, C12P21/02, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, PC
G01N33/53,
PC G01N33/566, C12N15/00, C12N5/00
CC Designed oligonucleotide primer to amplify DNA encoding TGR8
FH Key Location/Qualifiers
FT Source 1. 32
Location/Qualifiers

FEATURES

Source

1. 32
/organism="Synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 81.5%; Score 22; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 CATGAGCACACGACGCCAC 27
Db 6 CATGAGCACACGACGCCAC 27

RESULT 13

BD103849 32 bp DNA linear PAT 27-AUG-2002

LOCUS BD103849 Novel G protein-coupled receptor and its DNA.

DEFINITION BD103849

ACCESSION BD103849.1 GI:22649423

VERSION WO 0194582-A/1.

KEYWORDS WO 0194582-A/1.

SOURCE Synthetic construct

ORGANISM Artificial sequences.

REFERENCE 1 (bases 1 to 32)

AUTHORS Terao, Y., Matsui, H. and Shintani, Y.

TITLE Novel G protein-coupled receptor and its DNA

JOURNAL Patent: WO 0194582-A 1 13-DEC-2001;

COMMENT TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI

SHINTANI

OS Artificial Sequence

PN WO 0194582-A/1

PD 13-DEC-2001

PR 01-JUN-2001 WO 2001JP004643

PI 02-JUN-2000 JP 00P 170446, 23-JUN-2000 JP 00P 194926

PC YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI

PC C12N15/12, C12P21/02, C07K14/705, C07K16/28, A61K45/00, A61P25/00,

PC A61P29/00,

PC A61P9/00, A61P35/00, A61P37/00, A61P1/00, G01N33/50, G01N33/50,

PC G01N33/15

CC Designed oligonucleotide primer to amplify DNA encoding TGR8

FH Key Location/Qualifiers

FT Source 1. 32

Location/Qualifiers

1. 32
/organism="Artificial Sequence".

/organism="Synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

Query Match 81.5%; Score 22; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 CATGAGCACACGACGCCAC 27
Db 6 CATGAGCACACGACGCCAC 27

RESULT 14

AX463227 1059 bp DNA linear PAT 15-JUL-2002

LOCUS AX463227

DEFINITION Sequence 1 from Patent WO0248358.
ACCESSION AX463227
VERSION AX463227.1 GI:21886199
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE Smolyar, A., Zhu, Z., Encinas, J., Watanabe, S. and Okigami, H.
AUTHORS Regulation of human chemokine-like receptor
TITLE Patent: WO 0248358-A 1 20-JUN-2002;
JOURNAL Bayer Aktiengesellschaft (DE)
FEATURES
source 1..1059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 77.8%; Score 21; DB 6; Length 1059;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 ATGAGCACACGCGCCGAC 27
|||||
1 ATGAGCACACGCGCCGAC 21
Db 1 ATGAGCACACGCGCCGAC 21
|||||
RESULT 15
BD183203 1062 bp DNA linear PAT 17-JUN-2003
LOCUS BD183203
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD183203
VERSION BD183203.1 GI:31875403
KEYWORDS JP 2002345481-A/3.
SOURCE JP 2002345481-A/3.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1062)
REFERENCE Terao, Y., Matsui, H. and Shintani, Y.
AUTHORS Novel G protein-coupled receptor and its DNA
TITLE Patent: JP 2002345481-A 3 03-DEC-2002;
JOURNAL TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2002345481-A/3
PD 03-DEC-2002
PF 01-JUN-2001 JP 2001166688
PI YASUKO TERAO, HIDEKI MATSUJI, YASUSHI SHINTANI
PC C12N15/09, A61K45/00, A61P1/00, A61P3/00, A61P9/00, A61P25/00, PC
A61P29/00, PC
PC A61P35/00, A61P37/00, C07K14/705, C07K16/28, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, PC
G01N33/53,
PC G01N33/566, C12N15/00, C12N5/00
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
/organism="Homo sapiens (human)".
Location/Qualifiers
1..1062
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 77.8%; Score 21; DB 6; Length 1062;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 ATGAGCACACGCGCCGAC 27

Db 1 ATGAGCACACGCGCCGAC 21
|||||

Search completed: December 15, 2004, 00:42:54
Job time : 92.94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:00:04 ; Search time 12.6451 Seconds
(without alignments)
11208.653 Million cell updates/sec

Title: US-09-995-225B-41

Perfect score: 27
Sequence: 1 ggcgtcatgagacacgcacgcacacac 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: N_Geneseq_23Sep04:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2001bs:*
7: geneseqn2002as:*
8: geneseqn2002bs:*
9: geneseqn2003as:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	27	ABT04896	ADT04896 Human G P
2	27	100.0	1130	ADT29672	ADT29672 Human G-P
3	27	100.0	1202	ABV73364	ABV73364 Human TGR
4	27	100.0	1526	ADT46858	ADT46858 Human TGR
5	27	100.0	1526	ACA60889	ACA60889 Human CDN
6	27	100.0	1526	ABT57024	ABT57024 cDNA enco
7	27	100.0	2189	ADT18120	ADT18120 Human G-P
8	27	100.0	2198	ACF05275	ACF05275 Human G-P
9	27	100.0	2273	ADT28955	ADT28955 Human nov
10	22	81.5	32	ABK15560	ABK15560 Novel G-P
11	22	81.5	42	ADT87864	ADT87864 G-coupled
12	22	81.5	42	ADT87864	ADT87864 G-coupled
13	21	77.8	25	ADT30468	ADT30468 Human NOV
14	21	77.8	28	ACCT1787	ACCT1787 Human nov
15	21	77.8	116	ACCT1795	ACCT1795 Human G P
16	21	77.8	1038	ABV73373	ABV73373 Mouse TGR
17	21	77.8	1038	ADT28957	ADT28957 Mouse nov
18	21	77.8	1059	ABN84269	ABN84269 Human che
19	21	77.8	1062	ABK15562	ABK15562 cDNA enco
20	21	77.8	1062	ABK15563	ABK15563 cDNA enco
21	21	77.8	1062	ABT04873	ABT04873 Human G P

22	21	77.8	1062	6	ABN84273	ABN84273 Human che
23	21	77.8	1062	10	ACC44115	ACC44115 Human AXO
24	21	77.8	1062	10	ACC71785	ACC71785 Human G P
25	21	77.8	1110	9	ACF05276	ACF05276 Human G-P
26	21	77.8	1119	4	AA515730	AA515730 DNA encod
27	21	77.8	1119	10	ADT87766	ADT87766 G-coupled
28	21	77.8	1119	12	ADT79322	ADT79322 NOV10a co
29	21	77.8	1119	12	ADT56001	ADT56001 DNA encod
30	21	77.8	1343	4	AA515731	AA515731 DNA encod
31	21	77.8	1343	10	ADT87768	ADT87768 G-coupled
32	21	77.8	1343	12	ADT79324	ADT79324 NOV11 cod
33	21	77.8	1343	12	ADT56003	ADT56003 DNA encod
34	21	77.8	1343	12	ADT70587	ADT70587 Orphan re
35	20.2	74.8	1398	10	ADG33813	ADG33813 Actinomyc
36	19.8	73.3	3796	4	AAH62741	AAH62741 Shrimp wh
37	19.8	73.3	305107	4	AAH62689	AAH62689 Shrimp wh
38	19.6	72.6	178	10	ADT81026	ADT81026 INSP010 m
39	19.6	72.6	8676	10	ADT81128	ADT81128 INSP010 m
40	19.6	72.6	8742	10	ADT81122	ADT81122 INSP010 n
41	19.6	72.6	8838	6	ABV93367	ABV93367 Human NOV
42	19.6	72.6	8862	10	ADT81136	ADT81136 INSP010 P
43	19.6	72.6	34980	6	ABO81849	ABO81849 Bitidobac
44	19.6	72.6	34980	6	ABO81848	ABO81848 Bitidobac
45	19	70.4	612	4	AAH67672	AAH67672 Human HTL

ALIGNMENTS

RESULT 1
ABT04896
ID ABT04896 standard; DNA; 27 BP.
AC ABT04896;
XX
DT 11-OCT-2002 (first entry)
XX
DE Human G protein coupled receptor hrup35 PCR primer SEQ ID NO: 41.
XX
KW Human; G-protein coupled receptor; GPCR; hrup28; hrup29; hrup30; hrup31;
XX hrup32; hrup33; hrup34; hrup35; hrup36; hrup37; PCR; primer; ss.
XX
OS Homo sapiens.
XX
PN WO200242461-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001WO-US044386.
XX
PR 27-NOV-2000; 2000US-0253404P.
XX
PR 12-DEC-2000; 2000US-0255366P.
XX
PR 20-FEB-2001; 2001US-0270286P.
XX
PR 20-FEB-2001; 2001US-0270286P.
XX
PR 06-APR-2001; 2001US-0282356P.
XX
PR 06-APR-2001; 2001US-0282356P.
XX
PR 06-APR-2001; 2001US-0282358P.
XX
PR 06-APR-2001; 2001US-0282365P.
XX
PR 14-MAY-2001; 2001US-0290917P.
XX
PR 31-JUL-2001; 2001US-0309208P.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX
DR WPI; 2002-56565/60.
XX
PT Novel endogenous and non-endogenous versions of G protein-coupled
XX receptor useful for identification of candidate compounds as receptor
XX agonists or antagonists for use as therapeutic agents.
XX
PS Example 1; Page 28; 84pp; English.
XX

CC The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a PCR primer used to isolate a GPCR coding sequence of the invention
XX
SQ Sequence 27 BP; 6 A; 12 C; 7 G; 2 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 27; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GCGCTCATGTGAGCACACGACGCCAC 27
1 |||||
Db 1 GCGCTCATGTGAGCACACGACGCCAC 27
XX
RESULT 2
AAD29672 standard; cDNA; 1130 BP.
XX
AC AAD29672;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human G-protein coupled receptor (GCRC-6) cDNA.
XX
XX Human; G-protein coupled receptor; GCRC-6; cell proliferative disorder;
XX neurologic; cardiovascular; gastrointestinal; autoimmune; inflammatory;
XX metabolic; hepatic; psoriasis; cancer; epilepsy; Alzheimer's disease;
XX Pick's disease; Huntington's disease; Parkinson's disease; hypertension;
XX atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;
XX osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
XX anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;
XX transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;
XX neurotropic; neuroprotective; cardiant; immunosuppressive; anorectic;
XX virocidic; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 12..1130
FT /*tag= a
FT /product= "Human GCRC-6"
FT sig_peptide 12..152
FT /*tag= b
FT mat_peptide 153..1127
FT /*tag= c
FT /product= "Human mature GCRC-6"
XX
XX MO200210387-A2.
XX
XX
XX 07-FEB-2002.
XX
XX
XX 25-JUL-2001; 2001MO-US023433.
XX
XX
XX 27-JUL-2000; 2000US-0221478P.
XX 03-AUG-2000; 2000US-0223268P.
XX 21-AUG-2000; 2000US-0227054P.
XX 08-SEP-2000; 2000US-0231121P.
XX 13-SEP-2000; 2000US-0232243P.
XX 15-SEP-2000; 2000US-0232691P.
XX 22-SEP-2000; 2000US-0235146P.
XX
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Thornton M, Patterson C, Ial P, Burford N, Yue H, Gandhi AR;
XX Elliot VS, Rankumar J, Kallik MB, Kallik DA, Wala NK, Hafalia Adu;
XX Yao MG, Lu Y, Tribouley CW, Policky JL, Kearney L, Graul RC;
XX Warren BA, Lee EA, Ding L;
XX
XX WPI; 2002-188744/24.
XX
XX P-PDB; AAE18645.
XX
DR

XX
XX New human G-protein coupled receptor polypeptide for diagnosis,
XX prevention and treatment of cell proliferative, neurological,
XX cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic
XX disorders.
XX
XX Claim 5; Page 141; 150pp; English.
XX
XX The invention relates to novel human G-protein coupled receptors (GCRC)
XX and their encoding polynucleotides. GCRC is useful as an immunogen for
XX preparing monoclonal and polyclonal antibodies. GCRC is useful for
XX diagnosing, treating and preventing a cell proliferative disorder (e.g.,
XX hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy,
XX Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's
XX disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,
XX myocardial infarction), gastrointestinal disorder (e.g., gastritis,
XX cirrhosis, Crohn's disease), an autoimmune/inflammatory disorder (e.g.,
XX acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma,
XX rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,
XX osteoporosis), and viral infections. GCRC is useful in a number of drug
XX screening techniques, and to analyse the proteome of a tissue or cell
XX type. GCRC is useful for creating knockin humanised animals or
XX transgenic animals to model human diseases, in somatic or germline gene
XX therapy, to generate a transcript image of a tissue or cell type, for
XX detecting differences in the chromosomal location due to translocation,
XX inversion, etc., among normal, carrier or affected individuals, and as
XX hybridization probes for mapping naturally occurring genomic sequences.
XX GCRC is useful in Southern or Northern analysis, dot blot or other
XX membrane-based technologies, in PCR technologies, in dipstick, pin,
XX multiformat enzyme linked immunosorbant (ELISA)-like assays, and in
XX microarrays utilising fluids or tissues from patients to detect altered
XX GCRC expression. The present sequence is human GCRC-6 cDNA
XX
SQ Sequence 1130 BP; 258 A; 357 C; 226 G; 289 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 27; DB 6; Length 1130;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GCGCTCATGTGAGCACACGACGCCAC 27
1 |||||
Db 6 GCGCTCATGTGAGCACACGACGCCAC 32
XX
RESULT 3
ABV73364 standard; DNA; 1202 BP.
XX
XX
XX ABV73364;
XX
XX
XX 22-JAN-2003 (first entry)
XX
XX
XX Human TGR20 polypeptide encoding DNA.
XX
XX
XX G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory;
XX antileptic; neurotropic; neuroprotective; antianemic; antileuc; human;
XX antiparkinsonian; antilipemic; TGR20; gene; ds.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH 68..1129
FH CDS /*tag= a
FH FT /product= "TGR20"
FH FT /note= "GPCR polypeptide"
XX
XX
XX MO200277001-A2.
XX
XX
XX 03-OCT-2002.
XX
XX
XX 08-MAR-2002; 2002MO-US007171.
XX
XX
XX 09-MAR-2001; 2001US-00802803.
XX
XX
XX

PR 16-MAR-2001; 2001US-0276649P.
 XX
 XX (TULA-) TULARIK INC.
 XX
 XX Tian H, Zhao J, Chen J, Cutler G;
 DR WPI; 2003-018881/01.
 XX P-PSDB; ABB82499.
 PT New G-protein coupled receptor polypeptides and polynucleotides useful
 PT for identifying compounds for treating a TGR-associated disorder, e.g.
 PT psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's
 PT disease, anemia.
 XX
 PS Claim 6; Page 63; 87pp; English.
 XX
 XX The invention relates to G-protein coupled receptor (GPCR) polypeptides
 CC and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183,
 CC TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are
 CC useful for identifying compounds for treating a TGR-associated disorder,
 CC such as psoriasis, inflammatory bowel disease, hyperlipidemia,
 CC Parkinson's disease, Huntington's disease, anemia, immune and blood
 CC disorders, ulcerative colitis, Crohn's disease or spleen enlargement.
 CC They are also useful for identifying cells such as kidney, liver,
 CC hypothalamus, colon, adipose, or spleen cells, for forensics and
 CC paternity determination, diagnosing diseases and examining signal
 CC transduction. The present sequence represents a human TGR20 polypeptide
 CC encoding DNA
 XX
 SQ Sequence 1202 BP; 267 A; 380 C; 258 G; 297 T; 0 U; 0 Other;
 XX
 Query Match 100.0%; Score 27; DB 8; Length 1202;
 Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Gaps 0;
 Matches 27; Conservative 0; Indels 0; Indels 0; Gaps 0;
 QY 1 GCGCTCATGAGACACGACGCCCCAC 27
 Db 62 GCGCTCATGAGACACGACGCCCCAC 88
 RESULT 4
 ID AAD46858 standard; cDNA; 1526 BP.
 XX
 XX AAD46858;
 AC
 XX 27-JAN-2003 (first entry)
 DT
 XX
 DB Human 7TM domain receptor 65507 cDNA.
 XX
 XX Human; adenosine deaminase; seven transmembrane domain receptor; cancer;
 KW 7TM; glycoprotease; immune disorder; IGA deficiency; allergy; arrhythmia;
 KW rheumatoid arthritis; diabetes; atherosclerosis; cardiovascular disorder;
 KW hypertension; ischemic heart disease; obesity; myocardial infarction;
 KW endothelial cell disorder; Grave's disease; psoriasis; brain disorder;
 KW Parkinson's disease; Alzheimer's disease; haemopoietic disorder;
 KW cerebral oedema; metabolic disorder; liver disorder; platelet disorder;
 KW chromosome mapping; tissue typing; gene therapy; neuroprotective;
 KW cytoskeletal; anorectic; cardiant; haemostatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX
 PH Key Location/Qualifiers
 FT CDS 139..1200
 FT /tag= a
 FT /product= "Human 65507 protein"
 FT /note= "This region is specifically claimed as SEQ ID NO:
 FT 12 in claim 1 of the specification"
 FT sig_peptide 139..306
 FT /tag= b
 FT mat_peptide 307..1197
 FT /tag= c
 FT /product= "Human mature 65507 protein"

XX
 XX WO200274960-A2.
 XX
 XX
 PD 26-SEP-2002.
 XX
 XX 08-NOV-2001; 2001WO-US051427.
 XX
 XX 08-NOV-2000; 2000US-0246768P.
 PR 08-NOV-2000; 2000US-0246772P.
 PR 15-NOV-2000; 2000US-0249185P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX
 PI Leiby KR, Kapeller-Libermann R, Glucksmann M;
 XX
 XX WPI; 2002-759898/82.
 DR P-PSDB; AAE29236.
 XX
 XX New 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid molecule,
 PT useful for diagnosing and treating cancer, immune, cardiovascular,
 PT hematopoietic, brain, pain, metabolic, liver or platelet disorders, and
 PT in pharmacogenomics.
 XX
 XX
 PS Claim 1; Fig 17; 178pp; English.
 XX
 XX The present invention relates to novel 38650, 28472, 5495, 65507, 81588
 CC or 14354 nucleic acid molecules which encode adenosine deaminase, glyco-
 CC protease or seven transmembrane domain (7TM) receptor family members.
 CC Sequences of the invention are useful in diagnosing and treating cancer
 CC or aberrant cellular proliferation and/or differentiation (e.g. colon or
 CC lung cancer), immune disorders (e.g. selective IGA deficiency, rheumatoid
 CC arthritis, diabetes, allergy), heart or cardiovascular disorders (e.g.,
 CC hypertension, atherosclerosis, arrhythmias, ischemic heart disease,
 CC myocardial infarction, thrombus) including endothelial cell disorders
 CC (e.g. psoriasis, Grave's disease), haemopoietic disorders, brain
 CC disorders (e.g. cerebral oedema, Parkinson's or Alzheimer's disease),
 CC pain and metabolic disorders (e.g. obesity), liver disorders or platelet
 CC disorders. They are also useful in screening assays, predictive medicine
 CC (e.g. diagnostic assays, prognostic assays, monitoring clinical trials
 CC and pharmacogenetics) and prophylactic and therapeutic methods. The
 CC nucleic acids may also be used in chromosome mapping, tissue typing and
 CC forensic biology and as surrogate markers. Sequences of the invention are
 CC also used in gene therapy. The present sequence is human 7TM domain
 CC receptor 65507 cDNA
 XX
 SQ Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
 XX
 Query Match 100.0%; Score 27; DB 6; Length 1526;
 Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Gaps 0;
 Matches 27; Conservative 0; Indels 0; Indels 0; Gaps 0;
 QY 1 GCGCTCATGAGACACGACGCCCCAC 27
 Db 133 GCGCTCATGAGACACGACGCCCCAC 159
 RESULT 5
 ID ACA60889 standard; cDNA; 1526 BP.
 XX
 XX ACA60889;
 AC
 XX 08-UTL-2003 (first entry)
 DT
 XX
 DB Human cDNA encoding 65507 protein.
 XX
 XX Human; ss; gene; cancer; aberrant cellular proliferation;
 KW differentiation; immune disorders; heart disorder; brain disorder;
 KW cardiovascular disorder; endothelial cell disorder; pain disorder;
 KW haematopoietic disorder; blood vessel disorder; metabolic disorder;
 KW liver disorder; platelet disorder.
 XX
 OS Homo sapiens.

```

XX Key Location/Qualifiers
FH CDS 139..1200
FT /*tag= a
FT /product= "Protein 65507"
FT /note= "this CDS is specifically claimed in claim 1"
XX
XX US2003009017-A1.
XX
XX 09-JAN-2003.
XX
XX 08-NOV-2001; 2001US-00012140.
XX
XX 08-NOV-2000; 2000US-0246768P.
XX 08-NOV-2000; 2000US-0246772P.
XX 15-NOV-2000; 2000US-0249185P.
XX
XX (LEIB/) LEIBY K R.
XX (KAPE/) KAPPELLER-LIBERMANN R.
XX (GLUC/) GLUCKSMANN M A.
XX
XX LeiBy KR, Kapeller-libermann R, Glucksmann MA;
XX
XX WPI; 2003-428888/40.
XX P-PSDB; ABU09571.
XX
XX New isolated 38650, 28472, 5495, 65507, 81588 and 14354 nucleic acid
XX molecules, useful for diagnosing, treating cancer, pain, or immune,
XX heart, endothelial cell, hematopoietic, blood vessel, brain, metabolic
XX and liver disorders.
XX
XX Claim 2; Fig 17; 90pp; English.
XX
XX The invention relates to an isolated 38650 (encoding adenosine
XX deaminase), 28472 (encoding a glycoprotease), 5495 (encoding a 7
XX transmembrane receptor), 65507, 81588 or 14354 nucleic acid molecule (or
XX a sequence which is at least 60% identical to the six nucleic acids or
XX their open reading frames, fragments of at least 15 nucleotides,
XX naturally occurring variants, or a DNA insert of the plasmid deposited
XX with the American Type Culture Collection as Accession No. not defined in
XX the specification, which encodes the amino acid sequence). Also included
XX are a host cell containing the nucleic acids (used to produce the
XX proteins), the encoded proteins, an antibody that selectively binds to
XX the polypeptide, and identifying a compound that binds to/modulates the
XX activity of the polypeptide. The nucleic acid molecules, polypeptides and
XX methods are useful for diagnosing, treating cancer, aberrant cellular
XX proliferation and/or differentiation, immune disorders, heart disorders,
XX cardiovascular disorders including endothelial cell disorders,
XX haematopoietic disorders, blood vessel disorders, brain disorders, pain
XX and metabolic disorders, liver disorders and platelet disorders (many
XX examples of these disorders are given in the specification). The present
XX sequence is the Human cDNA 65507
XX
XX Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
XX
XX Query Match 100.0%; Score 27; DB 8; Length 1526;
XX Best Local Similarity 100.0%; Pred. No. 0.11;
XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GCGCTCATGGAGCAGCAGCAGCCAC 27
XX ||||||||||||||||||||
XX Db 133 GCGCTCATGGAGCAGCAGCAGCCAC 159
XX
XX RESULT 6
XX ID ABS57024 standard; cDNA; 1526 BP.
XX
XX ABS57024;
XX
XX 30-JAN-2003 (first entry)
XX
XX cDNA encoding novel human 7 transmembrane receptor 65507.

```

```

XX
XX Cancer; aberrant cell proliferation; aberrant cell differentiation;
XX breast cancer; ovarian cancer; prostate cancer; colon cancer;
XX lung cancer; immune disorder; heart disorder; cardiovascular disorder;
XX endothelial disorder; hematopoietic disorder; blood vessel disorder;
XX brain disorder; pain; metabolic disorder; liver disorder; diabetes;
XX platelet disorder; carcinoma; sarcoma; leukemia; Hodgkin's disease;
XX autoimmune disorder; hypertension; atherosclerosis; heart failure;
XX myocardial infarction; ischaemic heart disease; Crohn's disease;
XX Grave's disease; Kawasaki syndrome; Raynaud's disease; aneurysm;
XX cerebral ischaemia; peripheral neuropathy; Alzheimer's disease;
XX Parkinson's disease; anorexia nervosa; cachexia; gene; ss;
XX 7 transmembrane receptor.
XX
XX Homo sapiens.
XX
XX OS
XX
XX Key Location/Qualifiers
FH CDS 139..1200
FT /*tag= a
FT /product= "7 transmembrane receptor 65507"
FT /note= "Specifically claimed in claim 1"
XX
XX WO200277233-A2.
XX
XX 03-OCT-2002.
XX
XX 08-NOV-2001; 2001WO-US046724.
XX
XX 08-NOV-2000; 2000US-0246768P.
XX 08-NOV-2000; 2000US-0246772P.
XX 15-NOV-2000; 2000US-0249185P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX LeiBy KR, Kapeller-libermann R, Glucksmann M;
XX
XX WPI; 2003-029938/02.
XX P-PSDB; ABG71163.
XX
XX New adenosine deaminase, glycoprotease and seven transmembrane domain
XX nucleic acids and polypeptides, designated 38650, 28472, 5495, 65507,
XX 81588 and 14354, useful for treating e.g. leukemia, Hodgkin's disease or
XX hypertension.
XX
XX Claim 2; Fig 17; 178pp; English.
XX
XX The invention describes isolated 38650, 28472, 5495, 65507, 81588 and
XX 14354 nucleic acid molecules (I) and their encoded polypeptides (II). The
XX 38650 nucleic acid molecule comprises a sequence encoding adenosine
XX deaminase. The 28472 nucleic acid molecule comprises a sequence encoding
XX a human glycoprotease. The 5495, 65507, 81588 and 14354 comprise
XX sequences that encode a human seven transmembrane domain (7TM). The
XX 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid and polypeptide
XX sequences are useful for diagnosing, preventing or treating a subject
XX with or at risk of developing a disorder, e.g. cancer or aberrant
XX cellular proliferation and/or differentiation (e.g. breast, ovarian,
XX prostate, colon or lung cancer), immune disorders, heart disorders,
XX cardiovascular disorders, endothelial disorders, hematopoietic disorders,
XX blood vessel disorders, brain disorders, pain and metabolic disorders,
XX liver disorders or platelet disorders. These disorders include carcinoma,
XX sarcoma, leukemia, Hodgkin's disease, autoimmune disorders,
XX hypertension, atherosclerosis, heart failure, myocardial infarction,
XX ischaemic heart disease, Crohn's disease, Grave's disease, Kawasaki
XX syndrome, Raynaud's disease, aneurysm, cerebral ischaemia, peripheral
XX neuropathy, Alzheimer's disease, Parkinson's disease, anorexia nervosa,
XX cachexia or diabetes. This sequence encodes the novel human 7
XX transmembrane receptor 65507
XX
XX Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
XX
XX Query Match 100.0%; Score 27; DB 10; Length 1526;
XX Best Local Similarity 100.0%; Pred. No. 0.11;
XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Qy 1 GCGCTCATGAGCAGCAGCCGCAC 27
 |||||
 Db 133 GCGCTCATGAGCAGCAGCCGCAC 159

RESULT 7

ADD18120
 ID ADD18120 standard; DNA, 2189 BP.

AC ADD18120;

PT 15-JAN-2004 (first entry)

DE Human G-protein coupled receptor (GPCR) DNA sequence Seq ID6.

XX G protein coupled receptor; GPCR, signal transduction pathway; G protein;
 KM Alzheimer's disease; Parkinson's disease; diabetes; dwarfism;
 KM colour blindness; retinal pigmentosa; asthma; depression; schizophrenia;
 KM sleeplessness; hypertension; anxiety; stress; renal failure;
 KM cardiovascular disorder; neural disorder; oncology disorder;
 KM immune disorder; neuroprotective; gene therapy; gene; da.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 1047..2157
 /*tag= 8

FT /product= "Human G-protein coupled receptor"
 FT /trnal_except= (2085..2091, aa:KPIKS)

PN W02003016478-A2.

XX 27-FEB-2003.

PF 15-AUG-2002; 2002MO-US026017.

PR 20-SEP-2001; 2001US-0313658P.

PR 12-SEP-2001; 2001US-0318675P.

PR 30-OCT-2001; 2001US-0340703P.

PR 26-NOV-2001; 2001US-0333417P.

PR 06-DEC-2001; 2001US-0338367P.

PR 06-FEB-2002; 2002US-0355596P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Feder JN, Ramanathan CS, Gopal S, Mintier GA;

XX WPI; 2003-278558/27.

XX P-PSDB; ADD18023.

XX New nucleic acid, useful for manufacturing a medicament for preventing,

XX treating or ameliorating a medical condition e.g., neural disorder.

XX Example 2; SEQ ID NO 6; 251bp; English.

XX This invention relates to novel G protein coupled receptors (GPCRs) and
 CC their encoding nucleotide sequences. Many medically significant
 CC biological processes are mediated by proteins participating in signal
 CC transduction pathways involving G proteins. GPCRs are one of the largest
 CC receptor superfamilies known. These receptors are biologically important
 CC and malfunction of these receptors results in diseases such as
 CC Alzheimer's, Parkinson's, diabetes, dwarfism, colour blindness, retinal
 CC pigmentosa and asthma. They are also involved in depression,
 CC schizophrenia, sleeplessness, hypertension, anxiety, stress, renal
 CC failure and other cardiovascular, neural, oncology and immune disorders.
 CC A modulator of the GPCRs of the invention may have neuroprotective
 CC activity whilst the sequences of the invention may be useful for gene
 CC therapy. The invention may also be useful for manufacturing a medicament
 CC for preventing, treating or ameliorating a medical condition. The present
 CC sequence is the DNA sequence encoding a human GPCR of the invention.
 XX Sequence 2189 BP; 451 A; 691 C; 581 G; 466 T; 0 U; 0 Other;

Query Match 100.0%; Score 27; DB 10; Length 2189;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGCTCATGAGCAGCAGCCGCAC 27
 |||||
 Db 1041 GCGCTCATGAGCAGCAGCCGCAC 1067

RESULT 8

AC ACF05275
 ID ACF05275 standard; cDNA, 2198 BP.

AC ACF05275;

DT 06-NOV-2003 (first entry)

DE Human G-protein coupled receptor HGPBRMY34 cDNA.

XX HGPBRMY34; G-protein coupled receptor; receptor; GPCR-P14; GPCR-145;
 KM human; neuroprotective; nootropic; tranquilizer; antimalgraine;
 KM neuroleptic; antianemic; antidepressant; anticonvulsant; antiparkinsonian;
 KM cytostatic; cardiac; hypotensive; antidiabetic; analgesic; anorectic;
 KM anti-HIV; antisthmatic; osteopathic; uropathic; antitumor; antiallergic;
 KM gene therapy; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 1247..2165
 FT /*tag= a

FT /product= "HGPBRMY34"

PN W02003050256-A2.

XX 19-JUN-2003.

PD 06-DEC-2002; 2002MO-US039290.

PR 06-DEC-2001; 2001US-0338371P.

PR (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Feder JN, Gopal S, Mintier GA, Ramanathan CS;

XX WPI; 2003-577295/54.

XX P-PSDB; ABR62521.

XX New nucleic acid molecule encoding a human G-protein coupled receptor,
 PT HGPBRMY34, useful for diagnosing, preventing or treating diseases
 PT involving the receptor, for example Parkinson's disease, dementia,
 XX asthma, hypertension or cancer.

XX Claim 1; Fig 1A-B; 112bp; English.

XX The present sequence is that of cDNA encoding human HGPBRMY34, a newly
 CC identified G-protein coupled receptor (GPCR) belonging to the group of
 CC 'Class A' GPCRs and showing homology to the Pfam model 7 transmembrane
 CC receptor, rhodopsin family. The cDNA was isolated by screening expressed
 CC sequence tag databases using GPCR sequences. HGPBRMY34, also referred to
 CC as GPCR-P14 and/or GPCR-145, is highly expressed in brain (amygdala,
 CC caudate nucleus, corpus callosum, hippocampus, thalamus, substantia
 CC nigra), spinal cord and pituitary, indicating an association in
 CC neurological systems and conditions. It is also expressed in the bone
 CC marrow and testis. The invention provides HGPBRMY34 polynucleotides,
 CC polypeptides and antibodies, expression vectors, host cells and antisense
 CC molecules, methods for screening for modulators of HGPBRMY34 activity
 CC and/or function, and methods for diagnosing, treating, preventing and
 CC screening for disorders and diseases associated with abnormal HGPBRMY34
 CC activity, including: a disorder related to aberrant G-protein coupled
 CC signalling; a disorder related to aberrant cell cycle regulation;
 CC neurological disorders; anxiety; headache; migraine; schizophrenia; manic

CC depression; delirium; dementia; severe mental retardation and
 CC dyskinesias, such as Huntington's disease or Gilles de la Tourette's
 CC syndrome; Parkinson's disease; brain disorders; spinal cord disorders;
 CC affective disorders; neoplastic disorders; cardiovascular disorders;
 CC acute heart failure; hypotension; hypertension; angina pectoris;
 CC myocardial infarction; an immunological disorder; immune-related
 CC disorders; endocrine diseases; growth disorders; neuropathic pain;
 CC obesity; anorexia; HIV infections; cancers; bulimia; asthma; osteoporosis
 CC ; psychosis; metabolic disorders; pituitary disorders; urinary retention;
 CC ulcers; allergies; or benign prostatic hypertrophy (all claimed)
 XX
 SO Sequence 2198 BP; 456 A; 693 C; 561 G; 468 T; 0 U; 0 Other;
 Query Match 100.0%; Score 27; DB 9; Length 2198;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCGCTCATGGAGCACACGCCGCCAC 27
 DB 1041 GCGCTCATGGAGCACACGCCGCCAC 1067
 RESULT 9
 ID ADO28955 standard; cDNA; 2273 BP.
 AC ADO28955;
 DT 29-JUN-2004 (first entry)
 XX
 DE Human novel GPCR PCR3 polynucleotide, SEQ ID NO:54.
 XX
 KW G-protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW kidney disorder; metabolic disorder; nutritive disorder; cancer;
 KW joint disorder; uterus disorder; lung disorder; breast disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thyroid disorder; thyroid disorder; antiparkinsonian; antianemic;
 KW cytoskeletal; antinflammatory; vasotropic; antidiarrhoeic; antidiabetic;
 KW CNS; central nervous system; vasotropic; antidiarrhoeic; antidiabetic;
 KW vitruvian; hepatotropic; antibacterial; antianemic; antidiarrhoeic;
 KW dermatological; antitumor; antihypertensive; anorectic;
 KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2004040000-A2.
 XX
 PD 13-MAY-2004.
 XX
 PF 09-SEP-2003; 2003WO-US028226.
 XX
 PR 09-SEP-2002; 2002US-0409303P.
 XX
 PR 09-APR-2003; 2003US-0461329P.
 XX
 PA (PRIM-) PRIMAL INC.
 XX
 PI Galanaris GA, Bergmann JF, Gragerov A, Hohmann J, Li F,
 PI Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
 DR WPI; 2004-390329/36.
 DR P-PSDB; ADO28954.
 XX
 PT Novel mammalian G-protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX
 PS Claim 13, SEQ ID NO 54, 542bp; English.

XX
 CC The invention relates to human and mouse G-protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention, methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
 CC nucleic acid of the invention. Note: The full sequence data for this
 CC patent did not form part of the printed specification, those sequences
 CC not shown were obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences.
 XX
 SO Sequence 2273 BP; 503 A; 691 C; 549 G; 530 T; 0 U; 0 Other;
 Query Match 100.0%; Score 27; DB 12; Length 2273;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCGCTCATGGAGCACACGCCGCCAC 27
 DB 447 GCGCTCATGGAGCACACGCCGCCAC 473
 RESULT 10
 ID ABR15560 standard; DNA; 32 BP.
 AC ABR15560;
 DT 08-MAY-2002 (first entry)
 XX
 DE Novel G-protein coupled receptor TGR8, PCR primer TGR8sa1.
 XX
 KW G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
 KW cytoskeletal; antinflammatory; antitumor; fetal brain;
 KW central nervous system disease; circulatory organ disorder; cancer;
 KW metabolic disease; immunological disease; gastrointestinal disease;
 KW gene therapy; transgenic animal; human; PCR; primer; ss; TGR8sa1.
 XX
 OS Homo sapiens.
 XX
 PN WO200194582-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 01-JUN-2001; 2001WO-JP004643.
 XX
 PR 02-JUN-2000; 2000JP-00170446.
 XX
 PR 23-JUN-2000; 2000JP-00194926.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.

PI Terao Y, Matsui H, Shintani Y;
 XX
 DR WPI, 2002-164317/21.
 XX
 PT Human fetal brain-originated G protein-coupled receptor protein TGR8 and
 PT encoding DNA, for developing drugs to treat e.g. diseases of the central
 PT nervous system or circulatory organs, cancer, and metabolic diseases.
 XX
 PS Example 2; Page 82; 102pp; Japanese.
 XX
 CC The invention describes a human fetal brain-originated G protein-coupled
 CC receptor protein, or its salt. The protein and encoded DNA are useful for
 CC developing drugs to treat e.g. diseases of the central nervous system or
 CC circulatory organs, cancer, metabolic diseases, immunological diseases
 CC and gastrointestinal diseases. The invention also describes creation of a
 CC probe for gene therapy and construction of a transgenic animal. This
 CC sequence represents the PCR primer TGR8a1 used in the isolation of the
 CC novel human G-protein coupled receptor, TGR8
 XX
 SQ Sequence 32 BP; 7 A; 14 C; 8 G; 3 T; 0 U; 0 Other;
 Query Match 81.5%; Score 22; DB 6; Length 32;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 CATGAGACACACGACGCCAC 27
 DB 6 CATGAGACACACGACGCCAC 27
 RESULT 11
 ADJ87864
 ID ADJ87864 standard; DNA; 42 BP.
 AC ADJ87864;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE G-coupled protein receptor-related protein PCR primer #75.
 XX
 KW novel protein; G-coupled protein receptor-related protein;
 KW cardiomyopathy; atherosclerosis; cell signal processing-related disorder;
 KW metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
 KW Huntington's disease; epilepsy; anxiety; pain; hypercholesterolaemia;
 KW obesity; hypertension; Crohn's disease; systemic lupus erythematosus;
 KW viral infections; bacterial infection; parasitic infection;
 KW hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;
 KW Alzheimer's disease; tuberculous sclerosis; hypercalcaemia; cerebral palsy;
 KW PCR; ss; primer.
 XX
 OS Unidentified.
 XX
 PN WO2002102321-A2.
 XX
 PD 27-DEC-2002.
 XX
 PF 18-JUN-2002; 2002WO-US019522.
 XX
 PR 18-JUN-2001; 2001US-029894P.
 PR 18-JUN-2001; 2001US-0299134P.
 PR 04-OCT-2001; 2001US-00972446.
 PR 06-JUN-2002; 2002US-00299134.
 PR 07-JUN-2002; 2002US-00298994.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Anderson DM, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;
 PI Pena CE, Spaderna SK, Zhong M;
 XX
 DR WPI, 2003-167441/16.
 XX
 PT New MOLx polypeptides and polymucleotides, useful in gene therapy,
 PT particularly for treating or preventing e.g. cardiomyopathy,

PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
 PT cancer, stroke or pain.
 XX
 PS Example 10; SEQ ID NO 298; 378pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of novel G-
 CC coupled protein receptor-related (MOL) proteins. The DNA and protein
 CC sequences of the invention are useful for treating or preventing a MOL-
 CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
 CC associated with cell signal processing and metabolic pathway modulation,
 CC or diabetes. The DNA and protein sequences are also useful for the
 CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
 CC pain, hypercholesterolaemia, obesity, hypertension, Crohn's disease,
 CC systemic lupus erythematosus, viral infections, bacterial infections,
 CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
 CC syndrome, Alzheimer's disease, tuberculous sclerosis, hypercalcaemia, or
 CC cerebral palsy. The present DNA sequence represents a PCR primer that was
 CC used in the exemplification of the invention.
 XX
 SQ Sequence 42 BP; 9 A; 20 C; 10 G; 3 T; 0 U; 0 Other;
 Query Match 81.5%; Score 22; DB 10; Length 42;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 CATGAGACACACGACGCCAC 27
 DB 12 CATGAGACACACGACGCCAC 33
 RESULT 12
 ADO56045
 ID ADO56045 standard; DNA; 42 BP.
 AC ADO56045;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Human NOV 11, PCR primer #17.
 XX
 KW human; cancer; obesity; diabetic nephropathy; acute pancreatitis; stroke;
 KW multiple sclerosis; ss; primer; PCR.
 XX
 OS Homo sapiens.
 XX
 PN US2004058862-A1.
 XX
 PD 25-MAR-2004.
 XX
 PF 18-SEP-2002; 2002US-00246583.
 XX
 PR 18-SEP-2002; 2002US-00246583.
 XX
 PA (MAJU/) MAJUMDER K.
 XX
 PI Majumder K;
 XX
 DR WPI, 2004-268835/25.
 XX
 PT Novel NOVx polypeptides, useful for treating cancer, obesity, diabetic
 PT nephropathy, acute pancreatitis, strokes and multiple sclerosis.
 XX
 PS Example 9; Page 83; 87pp; English.
 XX
 CC The invention relates to novel isolated NOVx nucleic acids and encoded
 CC polypeptides. The nucleic acids, polypeptides and antibodies raised
 CC against the polypeptides are useful for preventing or treating diseases
 CC associated with aberrant NOVx expression or activity e.g., cancer,
 CC obesity, diabetic nephropathy, acute pancreatitis, strokes, multiple
 CC sclerosis. The present sequence represents a PCR primer used to clone
 CC NOVx nucleic acids of the invention.
 XX
 SQ Sequence 42 BP; 9 A; 20 C; 10 G; 3 T; 0 U; 0 Other;

Query Match 81.5%; Score 22; DB 12; Length 42;
 Best Local Similarity 100.0%; Pred. NO. 11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CATGAGCAGCAGCAGCCAC 27
 12 CATGAGCAGCAGCAGCCAC 33

Db

RESULT 13
 ADO30468
 ID ADO30468 standard; DNA; 25 BP.
 XX ADO30468;
 AC 29-JUL-2004 (first entry)
 DT 29-JUL-2004 (first entry)
 DE Human novel GPCR PGR3 RT-PCR primer, SEQ ID NO:1571.
 XX

G protein-coupled receptor; GPCR; drug screening; diagnosis;
 transgenic mouse; neurological disorder; adrenal gland disorder;
 colon disorder; intestinal disorder; cardiovascular disorder;
 muscular disorder; blood disorder; immune disorder; bone disorder;
 kidney disorder; metabolic disorder; nutritive disorder; cancer;
 ovary disorder; uterus disorder; prostate disorder; testis disorder;
 skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 thymus disorder; thyroid disorder; antiparkinsonian; antiinfective;
 cytostatic; antiinflammatory; vasotropic; antiangiinal; antiarhythmic;
 CNS; central nervous system; respiratory; antidiarrheic; antidiabetic;
 vituclide; hepatotropic; antibacterial; antianemic; antiseborrheic;
 dermatological; antitumor; antihypertensive; antineoplastic;
 immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 PGR3; reverse transcription-PCR; RT-PCR; primer; ss.
 XX

Homo sapiens.
 OS
 XX
 XX W02004040000-A2.
 PN 13-MAY-2004.
 PD 13-MAY-2004.
 XX
 XX 09-SEP-2003; 2003WO-US028226.
 PF 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX
 XX (PRIM-) PRIMAL INC.
 PA
 XX Galtanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F,
 PI Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
 XX
 XX WPI; 2004-390329/36.
 DR

Novel mammalian G protein coupled receptors, useful for identifying
 compounds that modulates diagnosing and treating disease condition
 associated with GPCR dysfunction e.g. autoimmune diseases, angina
 pectoris, Parkinson's disease.
 PT
 XX
 XX
 PS Disclosure; SEQ ID NO 1571; 542bp; English.

The invention relates to human and mouse G protein-coupled receptors
 (GPCRs) and nucleic acids encoding them. The invention also relates to
 sequences at least 90% identical to the GPCR proteins and nucleic acids
 associated with GPCRs of the invention; preventing or diagnosing diseases
 associated with GPCRs of the invention; methods of screening for
 compounds useful in the treatment of GPCR-related diseases; a transgenic
 mouse comprising a GPCR gene of the invention; a mouse comprising a
 mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 from the transgenic mice; kits comprising several mice, each of which has
 a mutation in a different GPCR gene of the invention, and kits comprising
 probes which hybridise to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors

comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 be used in the diagnosis, treatment or prevention of a wide variety of
 diseases including neurological disorders (e.g., Alzheimer's disease,
 depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 disorders of the adrenal gland; disorders of the colon or intestine
 (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 myocardial infarction); muscular disorders; blood disorders (e.g.,
 anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 diseases); and disorders of the kidney, liver, lung, breast, ovary,
 uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 thyroid (e.g., cancer). The present sequence represents a PCR primer
 used in the isolation of cDNA encoding the novel human GPCR PGR3. Note:
 CC The full sequence data for this patent did not form part of the printed
 CC specification; those sequences not shown were obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
 XX

Sequence 25 BP; 6 A; 11 C; 6 G; 2 T; 0 U; 0 Other;

QY 7 ATGAGCAGCAGCAGCCAC 27
 1 ATGAGCAGCAGCAGCCAC 21

Db

RESULT 14
 ACC71787
 ID ACC71787 standard; DNA; 28 BP.
 XX
 XX AC C71787;
 AC
 XX 25-JUL-2003 (first entry)
 DT
 XX
 XX Human G protein-coupled receptor related DNA fragment #SEQ ID 5.
 DE
 XX
 XX Human; G protein-coupled receptor; anorectic; eating disorder; obesity;
 KW
 XX ds.
 XX
 XX Homo sapiens.
 OS
 XX
 XX W02003027142-A1.
 PN 03-APR-2003.
 PD 03-APR-2003.
 XX
 XX 19-SEP-2002; 2002WO-JP009626.
 PF 21-SEP-2001; 2001JP-00288278.
 PR
 XX
 XX (YAMA) YAMANOUCHI PHARM CO LTD.
 PA
 XX Matsumoto S, Takasaki J, Kurama T, Saito T, Kamohara M, Soga T,
 PI Hiayama H;
 XX
 XX WPI; 2003-333291/31.
 DR

New G protein-coupled receptor and encoded gene, useful in screening
 PT preventives or remedies for eating disorders or obesity.
 XX
 XX Example 1; Page 39; 48bp; Japanese.
 PS

The invention relates to a novel G protein-coupled receptor. The protein
 CC and its encoded gene are useful for screening preventives or remedies for
 CC eating disorders or obesity. The current sequence represents a G protein
 CC coupled receptor related DNA fragment
 XX
 XX Sequence 28 BP; 7 A; 12 C; 7 G; 2 T; 0 U; 0 Other;

Query Match 77.8%; Score 21; DB 10; Length 28;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCAGCAGCAGCGCCAC 27
 |||||
 1 ATGAGCAGCAGCAGCGCCAC 21

RESULT 15

ACC71795

ID ACC71795 standard; DNA; 116 BP.

AC ACC71795;

DT 25-JUL-2003 (first entry)

DE Human G protein-coupled receptor related PCR primer #SEQ ID 13.

XX Human; G protein-coupled receptor; anorectic; eating disorder; obesity;

KM PCR; primer; ss.

XX Homo sapiens.

XX WO2003027142-A1.

PD 03-APR-2003.

PF 19-SEP-2002; 2002WO-JP009626.

PR 21-SEP-2001; 2001JP-00288278.

PA (YAMA) YAMANOUCHI PHARM CO LTD.

PI Matsumoto S, Takasaki J, Kurama T, Saito T, Kamohara M, Soga T;

XX Hiyyama H;

DR WPI; 2003-333291/31.

XX New G protein-coupled receptor and encoded gene, useful in screening

PT preventives or remedies for eating disorders or obesity.

XX Example 5; Page 41; 48pp; Japanese.

CC The invention relates to a novel G protein-coupled receptor. The protein

CC and its encoded gene are useful for screening preventives or remedies for

CC eating disorders or obesity. The current sequence represents a G protein

CC coupled receptor related PCR primer sequence

SQ Sequence 116 BP; 32 A; 34 C; 27 G; 23 T; 0 U; 0 Other;

QY 7 ATGAGCAGCAGCAGCGCCAC 27
 |||||
 89 ATGAGCAGCAGCAGCGCCAC 109

Search completed: December 14, 2004, 22:43:55
 Job time : 13.6451 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 14, 2004, 22:27:19 ; Search time 2.27127 Seconds
(without alignments)
8449.603 Million cell updates/sec

Title: US-09-995-225B-41

Perfect score: 27
Sequence: 1 ggcgcacatggagcacacgcacgcacacac 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5C COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5D COMB.seq.*
5: /cgn2_6/ptodata/1/ina/5E COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	70.4	720	3	US-08-913-014A-4
2	19	70.4	720	4	US-09-653-285-4
3	19	70.4	810	3	US-09-072-893C-5
4	19	70.4	1017	4	US-09-027-287-38
5	19	70.4	1017	4	US-09-252-656B-38
6	19	70.4	1017	4	US-09-523-323-38
7	19	70.4	1169	4	US-09-027-287-1
8	19	70.4	1169	4	US-09-252-656B-1
9	19	70.4	1169	4	US-09-523-323-1
10	19	70.4	1353	3	US-08-913-014A-6
11	19	70.4	1353	4	US-09-653-285-6
12	19	70.4	1491	3	US-08-913-014A-5
13	19	70.4	1491	4	US-09-653-285-5
14	18.6	68.9	1533	4	US-09-270-767-12691
15	18.2	67.4	1458	4	US-10-011-200-1
16	18.2	67.4	2346	4	US-10-011-200-2
17	18	66.7	438	3	US-09-060-756-600
18	18	66.7	438	4	US-09-670-314-600
19	18	66.7	492	3	US-09-060-756-587
20	18	66.7	492	4	US-09-670-314-587
21	18	66.7	1423	1	US-08-386-729A-6
22	18	66.7	2259	4	US-09-489-039A-3024
23	18	66.7	4403765	3	US-09-103-840A-2
24	18	66.7	4403765	3	US-09-103-840A-1
25	18	66.7	4411529	3	US-09-103-840A-2
26	18	66.7	4411529	3	US-09-103-840A-1
27	17.6	65.2	909	4	US-08-913-816C-1

28	17.6	65.2	3588	1	US-08-197-792-32	Sequence 32, Appl
29	17.6	65.2	3588	1	US-08-459-850-32	Sequence 32, Appl
30	17.6	65.2	3588	1	US-08-459-214-32	Sequence 32, Appl
31	17.4	64.4	223	3	US-08-205-697A-14	Sequence 14, Appl
32	17.4	64.4	223	3	US-08-702-525-14	Sequence 14, Appl
33	17.4	64.4	223	3	US-09-837-867A-14	Sequence 14, Appl
34	17.4	64.4	223	5	PCT-US95-02576-14	Sequence 14, Appl
35	17.4	64.4	393	4	US-09-252-991A-9425	Sequence 9425, Ap
36	17.4	64.4	453	4	US-09-252-991A-10826	Sequence 10826, A
37	17.4	64.4	491	4	US-09-252-991A-9486	Sequence 9486, Ap
38	17.4	64.4	957	4	US-09-252-991A-6986	Sequence 6986, Ap
39	17.4	64.4	1081	3	US-09-372-422A-33	Sequence 7011, Ap
40	17.4	64.4	1125	4	US-09-252-991A-7011	Sequence 7011, Ap
41	17.4	64.4	1188	4	US-09-252-991A-5131	Sequence 5131, Ap
42	17.4	64.4	1261	3	US-08-205-697A-12	Sequence 12, Appl
43	17.4	64.4	1261	3	US-08-702-525-12	Sequence 12, Appl
44	17.4	64.4	1261	4	US-09-837-867A-12	Sequence 12, Appl
45	17.4	64.4	1261	5	PCT-US95-02576-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-913-014A-4/C
; Sequence 4, Application US/08913014A
; Patent No. 6235878
; GENERAL INFORMATION:
; APPLICANT: Nishii, Kazumori
; APPLICANT: Hikiuchi, Yukiko
; APPLICANT: Shintani, Yasushi
; TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS
; TITLE OF INVENTION: PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: David G. Conlin, Esq.
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & USHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston,
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913, 014A
; FILING DATE: 04-SEP-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP97/02480
; FILING DATE: July 17, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: David G. Conlin
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 342/47694
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 720 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA
; US-08-913-014A-4
Query Match 70.4%; Score 19; DB 3; Length 720;
Best Local Similarity 81.5%; Pred. No. 24;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGCTCATGAGCAGCAGCGCCGAC 27
Db 669 GCGTTCATCCAGCAGCAGCGACGAC 643

RESULT 2

US-09-653-285-4/C
Sequence 4, Application US/09653285
Patent No. 6590090

GENERAL INFORMATION:
APPLICANT: Nishi, Kazumori
Hikichi, Yukiko
Shintani, Yasushi

TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS
PRODUCTION AND USE

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESSES:
ADDRESSEE: David G. Conlin, Esq.
DIKE, BRONSTEIN, ROBERTS & USHMAN, LLP

STREET: 130 Water Street
CITY: Boston,
STATE: MA

COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/653, 285

FILING DATE: 31-Aug-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/913,014

FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: David G. Conlin
REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 342/47694
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440

TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 720 base pairs

TYPE: Nucleic acid
STRANDEDNESS: Double

TOPOLOGY: Linear
MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-653-285-4

Query Match 70.4%; Score 19; DB 4; Length 720;
Best Local Similarity 81.5%; Pred. No. 24;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGCTCATGAGCAGCAGCGCCGAC 27
Db 669 GCGTTCATCCAGCAGCAGCGACGAC 643

RESULT 3

US-09-072-993C-5/C
Sequence 5, Application US/09072993C
Patent No. 6346388

GENERAL INFORMATION:
APPLICANT: Michael R. Brigham-Burke

APPLICANT: Peter R. Young
TITLE OF INVENTION: A METHOD OF IDENTIFYING AGONIST AND

TITLE OF INVENTION: ANTAGONISTS FOR TUMOR NECROSIS RELATED RECEPTORS TR1 AND TR2
FILE REFERENCE: GH-50030
CURRENT APPLICATION NUMBER: US/09/072,993C
CURRENT FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/055,513
PRIOR FILING DATE: 1997-08-13
PRIOR APPLICATION NUMBER: 60/056,980
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/057,550
PRIOR FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 810
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-072-993C-5

Query Match 70.4%; Score 19; DB 3; Length 810;
Best Local Similarity 81.5%; Pred. No. 24;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGCTCATGAGCAGCAGCGCCGAC 27
Db 743 GCGTTCATCCAGCAGCAGCGACGAC 717

RESULT 4

US-09-027-287-38/C
Sequence 38, Application US/09027287A
Patent No. 6479254

GENERAL INFORMATION:
APPLICANT: Eber, Reinhard

APPLICANT: Yu, Guo-Liang
APPLICANT: Ruben, Steven M.

APPLICANT: Ullrich, Stephen
TITLE OF INVENTION: Apoptosis Inducing Molecule II

FILE REFERENCE: 1486.0650004
CURRENT APPLICATION NUMBER: US/09/027,287A

CURRENT FILING DATE: 1998-02-20
EARLIER APPLICATION NUMBER: US 09/003,886

EARLIER FILING DATE: 1998-01-07
EARLIER APPLICATION NUMBER: US 08/822,953

EARLIER FILING DATE: 1997-03-21
EARLIER APPLICATION NUMBER: US 60/030,157

EARLIER FILING DATE: 1996-10-31
EARLIER APPLICATION NUMBER: US 60/013,923

EARLIER FILING DATE: 1996-03-22
NUMBER OF SEQ ID NOS: 55

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 38

LENGTH: 1017
TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: CDS
LOCATION: (1)..(624)
US-09-027-287-38

Query Match 70.4%; Score 19; DB 4; Length 1017;
Best Local Similarity 81.5%; Pred. No. 25;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGCTCATGAGCAGCAGCGCCGAC 27
Db 573 GCGTTCATCCAGCAGCAGCGACGAC 547

RESULT 5

US-09-252-656B-38/C
Sequence 38, Application US/09252656B
Patent No. 6495520

GENERAL INFORMATION:


```

; APPLICANT: Ebner, Reinhard
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M.
; APPLICANT: Zhang, Jun
; APPLICANT: Ullrich, Stephen
; APPLICANT: Zhai, Yifan
; TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
; FILE REFERENCE: 1488, 0650006
; CURRENT APPLICATION NUMBER: US/09/252,656B
; CURRENT FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: US 60/075,409
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 09/027,287
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 09/003,886
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: US 08/822,953
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/013,923
; PRIOR FILING DATE: 1996-03-22
; PRIOR APPLICATION NUMBER: US 60/030,157
; PRIOR FILING DATE: 1996-10-31
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(624)
; OTHER INFORMATION: CDNA
US-09-252-656B-38

```

```

Query Match          70.4%; Score 19; DB 4; Length 1017;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 GCGCTCATGGAGCAGCAGCGCCGAC 27
DB 573 GCGTTCATCCAGCAGCAGCGACGAC 547

```

```

RESULT 6
US-09-523-323-38/c
; Sequence 38, Application US/09523323
; Patent No. 6635743
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ullrich, Stephen
; APPLICANT: Zhai, Yifan
; TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
; FILE REFERENCE: 1488, 0650006
; CURRENT APPLICATION NUMBER: US/09/523,323
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 60/168,380
; EARLIER FILING DATE: 1999-12-02
; EARLIER APPLICATION NUMBER: 60/148,326
; EARLIER FILING DATE: 1999-08-11
; EARLIER APPLICATION NUMBER: 60/142,657
; EARLIER FILING DATE: 1999-07-06
; EARLIER APPLICATION NUMBER: 60/137,457
; EARLIER FILING DATE: 1999-06-04
; EARLIER APPLICATION NUMBER: 60/124,041
; EARLIER FILING DATE: 1999-03-11
; EARLIER APPLICATION NUMBER: 09/252,656
; EARLIER FILING DATE: 1999-02-19
; EARLIER APPLICATION NUMBER: 60/075,409
; EARLIER FILING DATE: 1998-02-20
; EARLIER APPLICATION NUMBER: 09/027,287
; EARLIER FILING DATE: 1998-02-20

```

```

; EARLIER APPLICATION NUMBER: 09/003,886
; EARLIER FILING DATE: 1998-01-07
; EARLIER APPLICATION NUMBER: 08/822,953
; EARLIER FILING DATE: 1997-03-21
; EARLIER APPLICATION NUMBER: 60/013,923
; EARLIER FILING DATE: 1996-03-22
; EARLIER APPLICATION NUMBER: 60/030,157
; EARLIER FILING DATE: 1996-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(624)
US-09-523-323-38

```

```

Query Match          70.4%; Score 19; DB 4; Length 1017;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 GCGCTCATGGAGCAGCAGCGCCGAC 27
DB 573 GCGTTCATCCAGCAGCAGCGACGAC 547

```

```

RESULT 7
US-09-027-287-1/c
; Sequence 1, Application US/09027287A
; Patent No. 6479254
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ullrich, Stephen
; TITLE OF INVENTION: Apoptosis Inducing Molecule II
; FILE REFERENCE: 1488, 0650004
; CURRENT APPLICATION NUMBER: US/09/027,287A
; CURRENT FILING DATE: 1998-02-20
; EARLIER APPLICATION NUMBER: US 09/003,886
; EARLIER FILING DATE: 1998-01-07
; EARLIER APPLICATION NUMBER: US 08/822,953
; EARLIER FILING DATE: 1997-03-21
; EARLIER APPLICATION NUMBER: US 60/030,157
; EARLIER FILING DATE: 1996-10-31
; EARLIER APPLICATION NUMBER: US 60/013,923
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (49)..(768)
US-09-027-287-1

```

```

Query Match          70.4%; Score 19; DB 4; Length 1169;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 GCGCTCATGGAGCAGCAGCGCCGAC 27
DB 717 GCGTTCATCCAGCAGCAGCGACGAC 691

```

```

RESULT 8
US-09-252-656B-1/c
; Sequence 1, Application US/09252656B
; Patent No. 6495520

```

```
/ GENERAL INFORMATION:
/ APPLICANT: Ebner, Reinhard
/ APPLICANT: Yu, Guo-liang
/ APPLICANT: Ruben, Steven M.
/ APPLICANT: Zhang, Jun
/ APPLICANT: Ulrich, Stephen
/ APPLICANT: Zhai, Yifan
/ TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
/ FILE REFERENCE: 1488.0650006
/ CURRENT APPLICATION NUMBER: US/09/252,656B
/ PRIOR FILING DATE: 1999-02-19
/ PRIOR APPLICATION NUMBER: US 60/075,409
/ PRIOR FILING DATE: 1998-02-20
/ PRIOR APPLICATION NUMBER: US 09/027,287
/ PRIOR FILING DATE: 1998-02-20
/ PRIOR APPLICATION NUMBER: US 09/003,886
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: US 08/822,953
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/013,923
/ PRIOR FILING DATE: 1996-03-22
/ PRIOR APPLICATION NUMBER: US 60/030,157
/ PRIOR FILING DATE: 1996-10-31
/ NUMBER OF SEQ ID NOS: 61
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1169
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (49)..(768)
/ OTHER INFORMATION: DNA (cDNA)
US-09-252-656B-1

Query Match          70.4%; Score 19; DB 4; Length 1169;
Best Local Similarity 81.5%; Pred. NO. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCGCCAC 27
Db 717 GCGTTCATCCAGCAGCAGCAGCCAC 691

RESULT 9
US-09-523-323-1/c
/ Sequence 1, Application US/09523323
/ Patent No. 6635743
/ GENERAL INFORMATION:
/ APPLICANT: Ebner, Reinhard
/ APPLICANT: Yu, Guo-liang
/ APPLICANT: Ruben, Steven M.
/ APPLICANT: Ulrich, Stephen
/ APPLICANT: Zhai, Yifan
/ TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
/ FILE REFERENCE: 1488.065000C
/ CURRENT APPLICATION NUMBER: US/09/523,323
/ PRIOR FILING DATE: 2000-03-10
/ PRIOR APPLICATION NUMBER: 60/168,380
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: 60/148,326
/ PRIOR FILING DATE: 1999-08-11
/ PRIOR APPLICATION NUMBER: 60/142,657
/ PRIOR FILING DATE: 1999-07-06
/ PRIOR APPLICATION NUMBER: 60/137,457
/ PRIOR FILING DATE: 1999-06-04
/ PRIOR APPLICATION NUMBER: 60/124,041
/ PRIOR FILING DATE: 1999-03-11
/ PRIOR APPLICATION NUMBER: 09/252,656
/ PRIOR FILING DATE: 1999-02-19
/ PRIOR APPLICATION NUMBER: 60/075,409
/ PRIOR FILING DATE: 1998-02-20
/ PRIOR APPLICATION NUMBER: 09/027,287
```

```
/ EARLIER FILING DATE: 1998-02-20
/ EARLIER APPLICATION NUMBER: 09/003,886
/ EARLIER FILING DATE: 1998-01-07
/ EARLIER APPLICATION NUMBER: 08/822,953
/ EARLIER FILING DATE: 1997-03-21
/ EARLIER APPLICATION NUMBER: 60/013,923
/ EARLIER FILING DATE: 1996-03-22
/ EARLIER APPLICATION NUMBER: 60/030,157
/ EARLIER FILING DATE: 1996-10-31
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 1169
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (49)..(768)
US-09-523-323-1

Query Match          70.4%; Score 19; DB 4; Length 1169;
Best Local Similarity 81.5%; Pred. NO. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCGCCAC 27
Db 717 GCGTTCATCCAGCAGCAGCAGCCAC 691

RESULT 10
US-08-913-014A-6/c
/ Sequence 6, Application US/08913014A
/ Patent No. 6235878
/ GENERAL INFORMATION:
/ APPLICANT: Nishi, Kazunori
/ APPLICANT: Hikichi, Yukiko
/ APPLICANT: Shintani, Yasushi
/ TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: David G. Conlin, Esq.
/ STREET: 130 Water Street
/ CITY: Boston,
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/913,014A
/ FILING DATE: 04-SEP-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP97/02480
/ FILING DATE: July 17, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: David G. Conlin
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 342/47694
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1353
/ TYPE: Nucleic acid
/ STRANDEDNESS: Double
```

TOPOLOGY: Linear
MOLECULE TYPE: CDNA
US-08-913-014A-6

Query Match 70.4%; Score 19; DB 3; Length 1353;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTATGAGCAGCAGCGCCAC 27
DB 743 GCGTTCATCCAGCAGCAGCAGCAGC 717

RESULT 11
US-09-653-285-6/C

Sequence 6, Application US/09653285
Patent No. 6590090
GENERAL INFORMATION:

APPLICANT: Nishi, Kazunori

Hikichi, Yukiko

Shintani, Yasushi

TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS

PRODUCTION AND USE

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: David G. Conlin, Esq.

DIKE, BRONSTEIN, ROBERTS & USHMAN, LLP

STREET: 130 Water Street

CITY: Boston,

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/653,285

FILING DATE: 31-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/913,014

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: David G. Conlin

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 342/47694

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1353

TYPE: Nucleic acid

STRANDEDNESS: Double

TOPOLOGY: Linear

MOLECULE TYPE: CDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-653-285-6

Query Match 70.4%; Score 19; DB 4; Length 1353;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

RESULT 12
US-08-913-014A-5/C

Sequence 5, Application US/08913014A
Patent No. 6235878
GENERAL INFORMATION:

APPLICANT: Nishi, Kazunori

Hikichi, Yukiko

Shintani, Yasushi

TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS

PRODUCTION AND USE

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: David G. Conlin, Esq.

DIKE, BRONSTEIN, ROBERTS & USHMAN, LLP

STREET: 130 Water Street

CITY: Boston,

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/913,014A

FILING DATE: 04-SEP-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP97/02480

FILING DATE: July 17, 1997

ATTORNEY/AGENT INFORMATION:

NAME: David G. Conlin

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 342/47694

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1491

TYPE: Nucleic acid

STRANDEDNESS: Double

TOPOLOGY: Linear

MOLECULE TYPE: CDNA

US-08-913-014A-5

Query Match 70.4%; Score 19; DB 3; Length 1491;
Best Local Similarity 81.5%; Pred. No. 26;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTATGAGCAGCAGCGCCAC 27
DB 1051 GCGTTCATCCAGCAGCAGCAGCAGC 1025

RESULT 13
US-09-653-285-5/C

Sequence 5, Application US/09653285
Patent No. 6590090
GENERAL INFORMATION:

APPLICANT: Nishi, Kazunori

Hikichi, Yukiko

Shintani, Yasushi

TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS

PRODUCTION AND USE

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: David G. Conlin, Esq.

DIKE, BRONSTEIN, ROBERTS & USHMAN, LLP

STREET: 130 Water Street

CITY: Boston,

STATE: MA

COUNTRY: USA

```
/
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/653,285
/ FILING DATE: 31-Aug-2000
/ CLASSIFICATION: <Unknown>
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/913,014
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: David G. Conlin
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 342/47694
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1491
/ TYPE: Nucleic acid
/ STRANDEDNESS: Double
/ TOPOLOGY: Linear
/ MOLECULE TYPE: cDNA
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-653-285-5

Query Match      70.4%; Score 19; DB 4; Length 1491;
Best Local Similarity 81.5%; Pred. No. 26;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 GCGCTCATGGACACGCGACGCCAC 27
DB      1051 GCGTTCATCCAGCAGCGACGACAC 1025

RESULT 14
US-09-270-767-12691/c
/ Sequence 12691, Application US/09270767
/ Patent No. 6703491
/
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 12691
/ LENGTH: 1533
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
/ FEATURE:
/ OTHER INFORMATION: n means any nucleotide
US-09-270-767-12691

Query Match      68.9%; Score 18.6; DB 4; Length 1533;
Best Local Similarity 84.0%; Pred. No. 38;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 GCGCTCATGGACACGCGACGCC 25
DB      734 GTGCTCATGGATCACACGACCC 710
```

```
RESULT 15
US-10-011-200-1
/ Sequence 1, Application US/10011200
/ Patent No. 6632631
```

```
/
/ GENERAL INFORMATION:
/ APPLICANT: Shuster, Jeffrey R
/ APPLICANT: Tanzer, Matthew M
/ APPLICANT: Hamer, Lisbeth
/ APPLICANT: Adachi, Kiyochi
/ APPLICANT: Dezaan, Todd M
/ APPLICANT: Lo, Sze Chung C
/ APPLICANT: Montenegro-Chamorro, Maria V
/ APPLICANT: Darveaux, Blaise A
/ APPLICANT: Frank, Sheryl A
/ APPLICANT: Heiniger, Ryan W
/ APPLICANT: Mahanty, Sanjoy K
/ APPLICANT: Pan, Huaqin
/ APPLICANT: Skatchunes, Amy R
/ APPLICANT: Tarpey, Rex
/ TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF INHIBITORS OF HOMOCITRATE SYNTH
/ FILE REFERENCE: 2123US
/ CURRENT APPLICATION NUMBER: US/10/011,200
/ CURRENT FILING DATE: 2001-11-09
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: Patent In version 3.1
/ SEQ ID NO 1
/ LENGTH: 1458
/ TYPE: DNA
/ ORGANISM: Magnaporthe grisea
US-10-011-200-1

Query Match      67.4%; Score 18.2; DB 4; Length 1458;
Best Local Similarity 87.0%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      4 CTCATGGAGCAGCAGCGCCCA 26
DB      484 CTCATGGAGCAGCTCGCAGGCCA 506

Search completed: December 15, 2004, 02:56:25
Job time : 10.2713 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:25:49 ; Search time 13.7935 Seconds

(without alignments)
10792.424 Million cell updates/sec

Title: US-09-995-225B-41

Perfect score: 27

Sequence: 1 gcgcacatgagacacacgcacccac 27

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US04_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US03_PUBCOMB.seq:*

6: /cgn2_6/ptodata/1/pubpna/US02_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US01_PUBCOMB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US00_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US04_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US03_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US02_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US01_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US00_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	27	9	US-09-995-225-41
2	27	100.0	27	10	US-09-995-225-41
3	27	100.0	1130	16	US-10-333-946-25
4	27	100.0	1202	14	US-10-094-417-1
5	27	100.0	1526	14	US-10-012-140-10
6	27	100.0	2117	17	US-10-779-104-1
7	27	100.0	2189	14	US-10-219-834-6
8	27	100.0	2198	15	US-10-314-076-1
9	22	81.5	32	16	US-10-296-294A-1
10	22	81.5	42	16	US-10-174-364-117
11	22	81.5	42	16	US-10-246-583-117
12	21	77.8	1038	14	US-10-094-417-19

13	21	77.8	1059	16	US-10-450-590-1	Sequence 1, Appl1
14	21	77.8	1062	9	US-09-995-225-15	Sequence 15, Appl1
15	21	77.8	1062	10	US-09-995-225-15	Sequence 15, Appl1
16	21	77.8	1062	14	US-10-012-140-12	Sequence 12, Appl1
17	21	77.8	1062	16	US-10-296-294A-3	Sequence 3, Appl1
18	21	77.8	1062	16	US-10-296-294A-4	Sequence 4, Appl1
19	21	77.8	1062	16	US-10-450-590-9	Sequence 9, Appl1
20	21	77.8	1110	15	US-10-314-076-3	Sequence 3, Appl1
21	21	77.8	1119	10	US-09-813-432-19	Sequence 19, Appl1
22	21	77.8	1119	15	US-10-174-364-19	Sequence 19, Appl1
23	21	77.8	1119	16	US-10-246-583-19	Sequence 19, Appl1
24	21	77.8	1119	17	US-10-689-832-19	Sequence 19, Appl1
25	21	77.8	1133	10	US-09-813-432-21	Sequence 21, Appl1
26	21	77.8	1133	15	US-10-174-364-21	Sequence 21, Appl1
27	21	77.8	1343	16	US-10-246-583-21	Sequence 21, Appl1
28	21	77.8	1343	17	US-10-689-832-21	Sequence 21, Appl1
29	20.6	76.3	1530	18	US-10-425-115-115292	Sequence 21, Appl1
30	20.2	74.8	1398	16	US-10-417-700A-72	Sequence 72, Appl1
31	19.6	72.6	8838	15	US-10-093-463-81	Sequence 81, Appl1
32	19.6	72.6	2256646	17	US-10-470-565-1	Sequence 1, Appl1
33	19	70.4	612	16	US-10-468-161-30	Sequence 30, Appl1
34	19	70.4	672	18	US-10-425-115-169941	Sequence 169941, Appl1
35	19	70.4	720	16	US-10-468-161-4	Sequence 4, Appl1
36	19	70.4	810	13	US-10-066-205-5	Sequence 5, Appl1
37	19	70.4	1017	9	US-09-027-287-38	Sequence 38, Appl1
38	19	70.4	1017	9	US-09-252-656B-38	Sequence 38, Appl1
39	19	70.4	1017	16	US-10-375-680-38	Sequence 38, Appl1
40	19	70.4	1169	9	US-09-027-287-1	Sequence 1, Appl1
41	19	70.4	1169	9	US-09-252-656B-1	Sequence 1, Appl1
42	19	70.4	1169	10	US-09-967-604-5	Sequence 5, Appl1
43	19	70.4	1169	15	US-10-272-411-12	Sequence 12, Appl1
44	19	70.4	1169	15	US-10-218-547-33	Sequence 33, Appl1
45	19	70.4	1169	15	US-10-272-328A-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1

US-09-995-225-41

Sequence 41, Application US/09995225

Publication No. US20020193584A1

GENERAL INFORMATION:

APPLICANT: Chen, Ruoping

APPLICANT: Chu, Zhi Liang

APPLICANT: Dang, Huang T.

APPLICANT: Lowitz, Kevin P.

APPLICANT: Priddy, Cameron

TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human

TITLE OF INVENTION: Receptors

FILE REFERENCE: AREN-0308

CURRENT APPLICATION NUMBER: US/09/995,225

CURRENT FILING DATE: 2001-11-26

PRIOR APPLICATION NUMBER: 09/170,496

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: PCT/US99/23938

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: 60/253,404

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/255,366

PRIOR FILING DATE: 2000-12-12

PRIOR APPLICATION NUMBER: 60/270,286

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,365

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/270,266

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,032

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,358

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,356

PRIOR FILING DATE: 2001-04-06

```

; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020193584A1e1 Sequence
US-09-995-225-41
```

```
Query Match          100.0%; Score 27; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GCGCTCATGGAGCAGCAGCGCCGAC 27
Db      1 GCGCTCATGGAGCAGCAGCGCCGAC 27
```

```
RESULT 2
US-09-995-225-41
; Sequence 41, Application US/09995225
; Publication No. US20030139588A9
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Chen, Ruoping
```

```
; APPLICANT: Chu, Zhi Liang
```

```
; APPLICANT: Dang, Huong T.
```

```
; APPLICANT: Lowitz, Kevin P.
```

```
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
```

```
; FILE REFERENCE: AREN-0308
```

```
; CURRENT APPLICATION NUMBER: US/09/995,225
```

```
; PRIOR FILING DATE: 2001-11-26
```

```
; PRIOR APPLICATION NUMBER: 09/170,496
```

```
; PRIOR FILING DATE: 1998-10-13
```

```
; PRIOR APPLICATION NUMBER: PCT/US99/23938
```

```
; PRIOR FILING DATE: 1998-10-13
```

```
; PRIOR APPLICATION NUMBER: 60/253,404
```

```
; PRIOR FILING DATE: 2000-11-27
```

```
; PRIOR APPLICATION NUMBER: 60/255,366
```

```
; PRIOR FILING DATE: 2000-12-12
```

```
; PRIOR APPLICATION NUMBER: 60/270,286
```

```
; PRIOR FILING DATE: 2001-02-20
```

```
; PRIOR APPLICATION NUMBER: 60/282,365
```

```
; PRIOR FILING DATE: 2001-04-06
```

```
; PRIOR APPLICATION NUMBER: 60/270,266
```

```
; PRIOR FILING DATE: 2001-02-20
```

```
; PRIOR APPLICATION NUMBER: 60/282,032
```

```
; PRIOR FILING DATE: 2001-04-06
```

```
; PRIOR APPLICATION NUMBER: 60/282,358
```

```
; PRIOR FILING DATE: 2001-04-06
```

```
; PRIOR APPLICATION NUMBER: 60/282,356
```

```
; PRIOR FILING DATE: 2001-04-06
```

```
; PRIOR APPLICATION NUMBER: 60/290,917
```

```
; PRIOR FILING DATE: 2001-05-14
```

```
; PRIOR APPLICATION NUMBER: 60/309,208
```

```
; PRIOR FILING DATE: 2001-07-31
```

```
; NUMBER OF SEQ ID NOS: 67
```

```
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 41
```

```
; LENGTH: 27
```

```
; TYPE: DNA
```

```
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
```

```
; OTHER INFORMATION: No. US20030139588A9e1 Sequence
```

```
US-09-995-225-41
```

```
Query Match          100.0%; Score 27; DB 10; Length 27;
```

```
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GCGCTCATGGAGCAGCAGCGCCGAC 27
Db      1 GCGCTCATGGAGCAGCAGCGCCGAC 27
```

```
RESULT 3
US-10-333-946-25
```

```
; Sequence 25, Application US/10333946
```

```
; Publication No. US20040023252A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.
```

```
; APPLICANT: ARVIZU, Chandra S.; TAL, Preeti G.
```

```
; APPLICANT: BURFORD, Neil; YUE, Henry
```

```
; APPLICANT: GANDHI, Ameena R.; ELIOTT, Vicki S.
```

```
; APPLICANT: RAMKUMAR, Jayalaxmi; BAUGHN, Mariah R.
```

```
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Nandinder K.
```

```
; APPLICANT: HAFALIA, April J.A.; YAO, Monique G.
```

```
; APPLICANT: LU, Yan; TRIBOUNLEY, Catherine M.
```

```
; APPLICANT: POLICKY, Jennifer L.; KEARNEY, Liam
```

```
; APPLICANT: GRAUL, Richard C.; WARREN, Bridget A.
```

```
; APPLICANT: LEE, Ernestine A.; DING, Li
```

```
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
```

```
; FILE REFERENCE: PI-0176 USN
```

```
; CURRENT APPLICATION NUMBER: US/10/333,946
```

```
; PRIOR FILING DATE: 2003-01-22
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/23433
```

```
; PRIOR FILING DATE: 2001-07-25
```

```
; PRIOR APPLICATION NUMBER: US 60/221,478
```

```
; PRIOR FILING DATE: 2000-07-27
```

```
; PRIOR APPLICATION NUMBER: US 60/223,268
```

```
; PRIOR FILING DATE: 2000-08-03
```

```
; PRIOR APPLICATION NUMBER: US 60/227,054
```

```
; PRIOR FILING DATE: 2000-08-21
```

```
; PRIOR APPLICATION NUMBER: US 60/231,121
```

```
; PRIOR FILING DATE: 2000-09-08
```

```
; PRIOR APPLICATION NUMBER: US 60/232,243
```

```
; PRIOR FILING DATE: 2000-09-13
```

```
; PRIOR APPLICATION NUMBER: US 60/232,691
```

```
; PRIOR FILING DATE: 2000-09-15
```

```
; PRIOR APPLICATION NUMBER: US 60/235,146
```

```
; PRIOR FILING DATE: 2000-09-22
```

```
; NUMBER OF SEQ ID NOS: 38
```

```
; SOFTWARE: PERL Program
```

```
; SEQ ID NO 25
```

```
; LENGTH: 1130
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
```

```
; FEATURE:
```

```
; NAME/KEY: misc feature
```

```
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7474767CB1
```

```
US-10-333-946-25
```

```
Query Match          100.0%; Score 27; DB 16; Length 1130;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GCGCTCATGGAGCAGCAGCGCCGAC 27
Db      6 GCGCTCATGGAGCAGCAGCGCCGAC 32
```

```
RESULT 4
```

```
US-10-094-417-1
```

```
; Sequence 1, Application US/10094417
```

```
; Publication No. US20030045685A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Tian, Hui
```

```
; APPLICANT: Zhao, Jiaqiang
```

```
; APPLICANT: Chen, Jin-Long
```

```
; APPLICANT: Cutler, Gene
```

```
APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030045685A1el Receptors
; FILE REFERENCE: 018781-008110US
; CURRENT APPLICATION NUMBER: US/10/094,417
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/802,803
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/276,649
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1202
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
; NAME/KEY: CDS
; LOCATION: (68)..(1129)
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
US-10-094-417-1
```

```
Query Match          100.0%; Score 27; DB 14; Length 1202;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCGCTCATGAGCAGCAGCGCCAC 27
DB 62 GCGCTCATGAGCAGCAGCGCCAC 88
```

```
RESULT 5
US-10-012-140-10
; Sequence 10, Application US/10012140
; Publication No. US2003009017A1
; GENERAL INFORMATION:
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
; TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
; FILE REFERENCE: 381552004900
; CURRENT APPLICATION NUMBER: US/10/012,140
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/246,768
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,772
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,185
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1526
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (139)...(1200)
US-10-012-140-10
```

```
Query Match          100.0%; Score 27; DB 14; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCGCTCATGAGCAGCAGCGCCAC 27
DB 133 GCGCTCATGAGCAGCAGCGCCAC 159
```

RESULT 6

```
US-10-779-104-1
; Sequence 1, Application US/10779104
; Publication No. US20040161799A1
; GENERAL INFORMATION:
; APPLICANT: Andrew J. Murphy
; APPLICANT: Susan Cro11-Kalish
; TITLE OF INVENTION: KOR3L1-like-Proteins and Methods of Modulating KOR3L-Mediated Acti
; FILE REFERENCE: REG 1000A
; CURRENT APPLICATION NUMBER: US/10/779,104
; CURRENT FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: 60/447,447
; PRIOR FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/495,577
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-779-104-1
```

```
Query Match          100.0%; Score 27; DB 17; Length 2117;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCGCTCATGAGCAGCAGCGCCAC 27
DB 274 GCGCTCATGAGCAGCAGCGCCAC 300
```

```
RESULT 7
US-10-219-834-6
; Sequence 6, Application US/10219834
; Publication No. US20030096751A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE TH
; FILE REFERENCE: D0191 NP
; CURRENT APPLICATION NUMBER: US/10/219,834
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/313,658
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/340,703
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US 60/318,675
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 60/355,596
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 60/333,417
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/338,367
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 2189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-219-834-6
```

```
Query Match          100.0%; Score 27; DB 14; Length 2189;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCGCTCATGAGCAGCAGCGCCAC 27
DB 1041 GCGCTCATGAGCAGCAGCGCCAC 1067
```

```
RESULT 8
US-10-314-076-1
; Sequence 1, Application US/10314076
```

```
/ Publication No. US20030152977A1
/ GENERAL INFORMATION:
/ APPLICANT: Bristol-Myers Squibb Company
/ TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMT34, AND VARIANTS A
/ FILE REFERENCE: D0197NP
/ CURRENT APPLICATION NUMBER: US/10/314,076
/ CURRENT FILING DATE: 2002-12-06
/ PRIOR APPLICATION NUMBER: U.S. 60/338,371
/ PRIOR FILING DATE: 2001-12-06
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 2198
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1047)..(2162)
/ OTHER INFORMATION:
US-10-314-076-1
```

```
Query Match          100.0%; Score 27; DB 15; Length 2198;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GCGCTCATGGACGACGACGCCAC 27
Db      1041 GCGCTCATGGACGACGACGCCAC 1067
```

```
RESULT 9
US-10-296-294A-1
/ Sequence 1, Application US/10296294A
/ Publication No. US20040029224A1
/ GENERAL INFORMATION:
/ APPLICANT: TERAO, Yasuko
/ APPLICANT: MATSUI, Hideki
/ APPLICANT: SHINTANI, Yasuhiro
/ TITLE OF INVENTION: No. US20040029224A1e1 G Protein-Coupled Receptor and its DNA
/ FILE REFERENCE: 2734 USOP
/ CURRENT APPLICATION NUMBER: US/10/296,294A
/ CURRENT FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: PCT/JP01/04643
/ PRIOR FILING DATE: 2001-06-01
/ PRIOR APPLICATION NUMBER: JP 2000-170446
/ PRIOR FILING DATE: 2000-06-02
/ PRIOR APPLICATION NUMBER: JP 2000-194926
/ PRIOR FILING DATE: 2000-06-23
/ NUMBER OF SEQ ID NOS: 8
/ SEQ ID NO 1
/ LENGTH: 32
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Designed oligonucleotide primer to amplify DNA encoding TGR8
US-10-296-294A-1
```

```
Query Match          81.5%; Score 22; DB 16; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 CATGGACGACGACGCCAC 27
Db      6 CATGGACGACGACGCCAC 27
```

```
RESULT 10
US-10-174-364-117
/ Sequence 117, Application US/10174364
/ Publication No. US20030216308A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson et al.
```

```
/ TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-729CIP2
/ CURRENT APPLICATION NUMBER: US/10/174,364
/ CURRENT FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,836
/ PRIOR FILING DATE: 2000-03-29
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 128
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 117
/ LENGTH: 42
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
US-10-174-364-117
```

```
Query Match          81.5%; Score 22; DB 15; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 CATGGACGACGACGCCAC 27
Db      12 CATGGACGACGACGCCAC 33
```

```
RESULT 11
US-10-246-583-117
/ Sequence 117, Application US/10246583
/ Publication No. US20040058862A1
/ GENERAL INFORMATION:
/ APPLICANT: Majumder
/ TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-729CIP2CON1
/ CURRENT APPLICATION NUMBER: US/10/246,583
/ CURRENT FILING DATE: 2002-12-06
/ PRIOR APPLICATION NUMBER: 10/174,364
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
```



```

; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 117
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
; OTHER INFORMATION: primer
US-10-246-583-117

Query Match      81.5%; Score 22; DB 16; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 CATGAGCACACGCGCCGAC 27
DB      12 CATGAGCACACGCGCCGAC 33

RESULT 12
US-10-094-417-19
; Sequence 19, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiayang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: NO. US20030045685A1el Receptors
; FILE REFERENCE: 018781-008110US
; CURRENT APPLICATION NUMBER: US/10/094,417
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/802,803
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/276,649
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1038
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR20
; NAME/KEY: CDS
; LOCATION: (1)..(1038)
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR20
US-10-094-417-19

Query Match      77.8%; Score 21; DB 14; Length 1038;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 ATGAGACACGACGCGCCGAC 27
DB      1 ATGAGACACGACGCGCCGAC 21

RESULT 13
US-10-450-590-1
; Sequence 1, Application US/10450590
; Publication No. US20040076985A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
; FILE REFERENCE: LIO316 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/450,590
```

```

; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/254,923
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: US 60/280,110
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/299,474
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 1
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-450-590-1

Query Match      77.8%; Score 21; DB 16; Length 1059;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 ATGAGACACGACGCGCCGAC 27
DB      1 ATGAGACACGACGCGCCGAC 21

RESULT 14
US-09-995-225-15
; Sequence 15, Application US/09995225
; Publication No. US20020193584A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Priddy, Cameron
; TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 15
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020193584A1el Sequence
US-09-995-225-15
```

Query Match 77.8%; Score 21; DB 9; Length 1062;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCAGCAGCGCCGAC 27
|||||
Db 1 ATGAGCAGCAGCGCCGAC 21

RESULT 15

US-09-995-225-15
; Sequence 15, Application US/09995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Priddy, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human C
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 15
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139588A9el Sequence
US-09-995-225-15

Query Match 77.8%; Score 21; DB 10; Length 1062;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCAGCAGCGCCGAC 27
|||||
Db 1 ATGAGCAGCAGCGCCGAC 21

Search completed: December 15, 2004, 02:53:12
Job time : 15.7935 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:17:59 : Search time 85.0321 Seconds
(without alignments)
11570.559 Million cell updates/sec

Title: US-09-995-225b-41

Perfect score: 27

Sequence: 1 gcgcacatgagcagcagcagccac 27

Scoring table: IDENTITY_NUC

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:
1: gb_esc1:
2: gb_esc2:
3: gb_esc3:
4: gb_esc4:
5: gb_esc5:
6: gb_esc6:
7: gb_esc7:
8: gb_esc8:
9: gb_esc9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	358	2	BB870915 BB870915
2	20.6	76.3	693	2	BF430732 OG04A11T3
3	20.6	76.3	771	8	CC340173 OGUA29TH
4	20.6	76.3	786	5	BU207814 603102460
5	20.6	76.3	825	5	BZ714949 PUFAD34TD
6	20.2	74.8	109	8	CC396746 PUFAD34TD
7	20.2	74.8	539	5	BO143379 fmbic.pk0
8	19.6	72.6	356	5	BY324213
9	19.6	72.6	412	6	CB388569 OSTF101D3
10	19.6	72.6	455	8	AQ557735 HS_2083.A
11	19.6	72.6	450	8	AQ177428 HS_2206.A
12	19.6	72.6	504	8	AZ446249 IM0242A10
13	19.6	72.6	599	2	AM057282 ca03c06.Y
14	19.6	72.6	617	6	CB469087 sn03.A04.
15	19.6	72.6	623	8	AZ999005 2M0286N12
16	19.6	72.6	656	8	BZ308052 1d35b05.b
17	19.6	72.6	680	9	CC714295 OGWED32TV
18	19.6	72.6	784	8	BZ551201 pacel-60
19	19.6	72.6	810	8	BZ359952 1d84d12.b
20	19.6	72.6	857	7	CN204077 10c4467.G
21	19.6	72.6	938	8	BZ570003 msh2_1120
22	19.6	72.6	951	9	CG093748 PUFED09TD
23	19.6	72.6	988	9	CG093746 PUFED09TB
24	19.6	72.6	1195	8	BZ579041 msh2_6143

25	19.6	72.6	1606	8	BZ556139 pacel-60
26	19.6	72.6	8523	9	AY405945 Homo sapi
27	19.6	71.9	443	9	AG185710 Pan trogl
28	19.2	71.1	343	9	CE794750 tigr-g88-
29	19.2	71.1	761	6	CA103713 SCZHR108
30	19.2	71.1	793	7	CO031516 EST809900
31	19.2	71.1	812	7	CO031470 EST809854
32	19.2	71.1	917	6	CA134571 SCJFRT106
33	19.2	71.1	200	7	CS59591 170004241
34	19.2	70.4	210	9	CG149622 PUFAD58TD
35	19.2	70.4	248	8	CC031961 3591_1_12
36	19.2	70.4	312	8	CC030711 3591_1_11
37	19.2	70.4	321	1	AJ434656 AJ434656
38	19.2	70.4	334	8	CC029795 3591_1_11
39	19.2	70.4	333	5	BP871389 BP871389
40	19.2	70.4	339	1	AV600125 AV600125
41	19.2	70.4	345	9	AG164506 Pan trogl
42	19.2	70.4	357	2	AW416075 51097 MAR
43	19.2	70.4	360	7	CF861995 pds001xH
44	19.2	70.4	378	6	C19478 C19478 Rice
45	19.2	70.4	396	8	CC030868 3591_1_11

ALIGNMENTS

RESULT 1
BB870915
LOCUS
DEFINITION
BB870915 RIKEN full-length enriched, 11 days embryo spinal cord Mus musculus cDNA clone G630025017 5', mRNA sequence.

ACCESSION
BB870915

VERSION
BB870915.1

KEYWORDS
GT-11717125

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus

REFERENCE
Mus musculus

AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.

1 (bases 1 to 358)

Akimura, T., Arai, T., Carrin, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hirose, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Komono, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Nunashima, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watabiki, A., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, url: http://genome-gsc.riken.jp/

Carrin, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibata, K., Itoh, M., Komono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Komono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carrin, P., Sugihara, Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers
1..358
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G630025017"
/issue_type="spinal cord"
/dev_stage="11 days embryo"
/clone_lib="RIKEN full-length enriched, 11 days embryo spinal cord"

ORIGIN

Query Match 100.0%; Score 27; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGACACGACGCCCCAC 27
|||||
Db 225 GCGCTCATGAGACACGACGCCCCAC 251

RESULT 2 693 bp mRNA linear EST 20-FEB-2001
BU430732
LOCUS 0604A113 OG Oryza sativa (indica cultivar-group) cDNA clone
DEFINITION 0604A11 5' similar to RIBULOSE BIPHOSPHATE CARBOXYLASE SMALL CHAIN
C PRECURSOR, mRNA sequence.
BU430732
ACCESSION BF430732.1 GI:11442833

VERSION BF430732
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoae; Oryza.
1 (bases 1 to 693)

REFERENCE
AUTHORS Cushman,J.C. and Komura,M.
TITLE A cDNA library from Oryza sativa cv. Pokkali leaves enriched for
JOURNAL salt-responsive genes
COMMENT Unpublished (2000)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

PCR primers
FORWARD: T7
BACKWARD: T3
Plate: OG04 row: A column: 11
Seq primer: T3
High quality sequence stop: 350
POLYA=No.

FEATURES

source

Location/Qualifiers
1..693
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="Pokkali"
/db_xref="taxon:39946"
/clone="OG04A11"
/issue_type="leaf"
/dev_stage="7 day-old (treated with 150 mM NaCl for 1, 2 and 7 days)."
/lab_host="Escherichia coli"
/clone_lib="OG"

ORIGIN

Query Match 76.3%; Score 20.6; DB 2; Length 693;
Best Local Similarity 85.2%; Pred. No. 3.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGACACGACGCCCCAC 27
|||||
Db 426 GCGCTCATGAGACACGACGCCCCAC 452

RESULT 3 771 bp DNA linear GSS 16-MAY-2003
CC340173
LOCUS OGUA629TH ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0392E09,
DEFINITION genomic survey sequence.
CC340173
ACCESSION CC340173.1 GI:30809579

VERSION CC340173
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 771)

REFERENCE
AUTHORS Whiteley,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Kohling,T.,
Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Contact: Cathy Whiteley
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteley@cigr.org
Seq primer: TR
Classes: sheared ends.

FEATURES

source

Location/Qualifiers
1..771
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0392E09"
/clone_lib="ZM 0.7.1.5-KB"
/note="Vector: pBCSK-1; Site_1: HincII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 76.3%; Score 20.6; DB 8; Length 771;
Best Local Similarity 85.2%; Pred. No. 3.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGACACGACGCCCCAC 27
|||||
Db 514 GAGCTGCTGAGACACGACGCCCCAC 540

RESULT 4 786 bp mRNA linear EST 25-NOV-2002
BU207814/c
LOCUS 603102460F1 CSROCHN03 Gallus gallus cDNA clone CHEST36916 5', mRNA
DEFINITION sequence.
BU207814
ACCESSION BU207814.1 GI:25377508

VERSION BU207814
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE

Phasiennae; Gallus.

1 (bases 1 to 786)

Boardman, P.E., Sant-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,

Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

PUBMED

COMMENT

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 016123560409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

Source

Location/Qualifiers

1..786

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="ChEST36916"

/tissue_type="whole embryo"

/dev_stage="20-21"

/lab_host="DH10B"

/clone_lib="CSECHN03"

/note="Organ: whole embryo; Vector: Bluescript II KS(+);

Site_1: EcoRI; Site_2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer. Using

methylation C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunt-ended, ligated to NotI adapters, digested with

EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

Bluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91:9228-9232 and Bonaldi et al., Genome Research 6

(1996): 791, except that a significantly longer

reannealing hybridization was used."

ORIGIN

Query Match

Best Local Similarity 76.3%; Score 20.6; DB 5; Length 786;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 GCGCTCATGAGCAGCAGCCGAC 27

111 GCGCTCATGAGCAGCAGCCGAC 85

111 GCGCTCATGAGCAGCAGCCGAC 85

RESULT 5

LOCUS

B2714949 825 bp DNA linear GSS 19-FEB-2003

DEFINITION PUPAD34TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBtra272F20,

genomic survey sequence.

ACCESSION B2714949

VERSION B2714949.1 GI:28435046

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 825)

White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,

Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and

Bennetzen, J.

Maize Genomics Consortium

Unpublished (2003)

CONTACT: Cathy Whitelaw

FEATURES

Source

Location/Qualifiers

1..825

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBtra272F20"

/clone_lib="ZM_0.6_1.0_KB"

/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high

COT selected genomic DNA library"

ORIGIN

Query Match

Best Local Similarity 76.3%; Score 20.6; DB 8; Length 825;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 GCGCTCATGAGCAGCAGCCGAC 27

632 GAGCTGCTGAGCAGCAGCCGAC 658

RESULT 6

LOCUS

CC396746 109 bp DNA linear GSS 19-MAY-2003

DEFINITION PUPHD11TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBtra429B22,

genomic survey sequence.

ACCESSION CC396746

VERSION CC396746.1 GI:30876836

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 109)

White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,

Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and

Bennetzen, J.

Maize Genomics Consortium

Unpublished (2003)

CONTACT: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..109

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBtra429B22"

/clone_lib="ZM_0.6_1.0_KB"

/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high

COT selected genomic DNA library"

ORIGIN

Query Match

Best Local Similarity 74.8%; Score 20.2; DB 8; Length 109;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

3 GCTCATGAGCAGCAGCCGAC 27

87 GCTTATGAGCAGCAGCCGAC 63

RESULT 7
LOCUS B0143379 539 bp mRNA linear EST 24-APR-2002
DEFINITION fhmhc.ph004.e14 Metarhizium anisopliae sf. acridum ARSEF 324
ACCESSION B0143379
VERSION B0143379.1 GI:20280438
KEYWORDS EST.
SOURCE Metarhizium anisopliae var. acridum
ORGANISM Metarhizium anisopliae var. acridum
REFERENCE Eukaryota, Fungi, Ascomycota, Pezizomycotina, Sordariomycetes,
Hypocreomycetidae, Hypocreales, Clavicipitaceae, mitosporic
Clavicipitaceae; Metarhizium.
AUTHORS 1 (bases 1 to 539)
Freimoser, F.M., Screen, S., Baga, S., Hu, G. and St. Leger, R.J.
TITLE EST analysis of genes expressed by two different insect pathogenic
fungi during optimized secretion of proteins
JOURNAL unpublished (2002)
COMMENT Contact: Freimoser F. M.
Department of Entomology
University of Maryland
4112 Plant Sciences Building, College Park, MD 20742, USA
Tel: 301 405 16 13
Fax: 301 314 92 90
Email: ff34@mail.umd.edu.
FEATURES Location/Qualifiers
source
1..539
/organism="Metarhizium anisopliae var. acridum"
/mol_type="mRNA"
/strain="ARSEF 324"
/db_xref="taxon:92637"
/clone_1lb="Metarhizium anisopliae sf. acridum ARSEF 324"
/notes="Vector: Unizap; Metarhizium anisopliae sf. acridum
was grown on insect cuticle and chitin for 24 hours. A
cDNA library was constructed in the unidirectional Lambda
vector Unizap."
ORIGIN
Query Match 74.8%; Score 20.2; DB 5; Length 539;
Best Local Similarity 88.0%; Pred. No. 5.5e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 3 GCTCATGAGACACGACGCCAC 27
|||||
Db 314 GCTCATGATGCCACGACGCCAC 290
|||||
RESULT 8
LOCUS BY324213 356 bp mRNA linear EST 11-DEC-2002
DEFINITION BY324213 RIKEN full-length enriched, synovial fibroblasts Mus
musculus cDNA clone L030023G11 5', mRNA sequence.
ACCESSION BY324213
VERSION BY324213.1 GI:26514636
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata; Euteleostomi;
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae; Mus.
1 (bases 1 to 356)
Okazaki, Y., Furuno, M., Kanakawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osaio, N., Saito, R., Suzuki, H., Yamakata, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragan, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gardbald, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Malais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Perce, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Reiss, B., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Secou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterson, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBLISHED 22354683
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Shuhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tsgami, M., Waki, K., Watanabe, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Vassilis Aidinis (Biomedical Sciences
Research Center "Al. Fleming" Institute of Immunology 14-16 Al.
Fleming street 16672 Vathi, Greece) whose assistance we gratefully
acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
FEATURES Location/Qualifiers
source
1..356
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="L030023G11"
/cell_type="synovial fibroblasts"
/clone_1lb="RIKEN full-length enriched, synovial
fibroblasts"
ORIGIN
Query Match 72.6%; Score 19.6; DB 5; Length 356;
Best Local Similarity 84.6%; Pred. No. 9.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GCGCTCATGAGACACGACGCCCA 26
|||||

Db 31 GCCGACGACGACGACGACGCCCA 6

RESULT 9
CB388569
LOCUS
DEFINITION CB388569 412 bp mRNA linear EST 15-MAY-2003
ACCESSION CB388569
VERSION CB388569.1 GI:30730279
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 412)
Reboul, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M.,
Amstrong, C.M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T.,
Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S.,
Eidreza, G.A., Jena, S., Chavet, R., Papsotiroupolou, V.,
Tollas, P.P., Pracek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H.,
Ducette-Stamm, L., Hill, D.E. and Vidal, M.
C. elegans ORFome version 1.1: experimental verification of the
genome annotation and resource for proteome-scale protein
expression
Nat. Genet. (2003) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFome
cloning project : Contact david_hill@dfci.harvard.edu or
marc_vidal@dfci.harvard.edu
POLYX-No.

FEATURES
Source
Location/Qualifiers
1..412
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_lib="AD-wrmcDNA"
/note="The AD-wrmcDNA library was generated with poly(A)+
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pPC86"

ORIGIN
Query Match 72.6%; Score 19.6; DB 6; Length 412;
Best Local Similarity 84.6%; Pred. No. 9.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGCTCATGGACGACGACGCCCA 26
Db 354 GCGCTCATGGACGACGACGCCCA 379

RESULT 10
A0557735/c
LOCUS
DEFINITION A0557735 455 bp DNA linear GSS 29-MAY-1999
HS_2083_A2_B12_T7C CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=2083 Col=24 Row=C, genomic survey
sequence.
ACCESSION A0557735
VERSION A0557735.1 GI:4917467
KEYWORDS
GSS.

SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 455)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.htbc.washington.edu>
Plate: 2083 row: C column: 24
Seq primer: T7
Class: BAC ends
High quality sequence stop: 455.

FEATURES
Source
Location/Qualifiers
1..455
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2083 Col=24 Row=C"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
E-Coli DH10B"

ORIGIN
Query Match 72.6%; Score 19.6; DB 8; Length 455;
Best Local Similarity 84.6%; Pred. No. 9.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CGCTCATGGACGACGACGCCCA 27
Db 323 CGCTCATGGACGACGCCCA 298

RESULT 11
A0177428
LOCUS
DEFINITION A0177428 490 bp DNA linear GSS 17-OCT-1998
HS_2206_A2_P06 MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=2206 Col=12 Row=K, genomic survey
sequence.
ACCESSION A0177428
VERSION A0177428.1 GI:3574795
KEYWORDS
SOURCE
ORGANISM
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 490)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2206 Row: K Column: 12
 Class: BAC ends
 High quality sequence stop: 490.

FEATURES
 source
 Location/Qualifiers
 1..490

/organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=2206 Col=12 Row=K"
 /sex="male"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBelBact11, BAC clones in E-Coli DH10B"

ORIGIN

Query Match 72.6%; Score 19.6; DB 8; Length 490;
 Best Local Similarity 84.6%; Pred. No. 9.9e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 2 CGCTATGAGCAGCAGCGCCGAC 27
 Db 428 CCCCATGAGGACACTCCTCCAC 453

RESULT 12
 A2446249 504 bp DNA linear GSS 04-OCT-2000
 LOCUS 1M0242A10 Mouse 10kb plasmid UUCGIM library Mus musculus genomic
 DEFINITION clone UUCGIM0242A10 R, genomic survey sequence.
 ACCESSION A2446249
 VERSION A2446249.1 GI:10596873
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS 1 (bases 1 to 504)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Irlam, H., Longacre, S., Mahmoud, S., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausern, A. and Wright, D., Weis, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0242 row: A column: 10
 Seq primer: CACACGGAACACGCTATGACC
 Class: Plasmid ends
 High quality sequence stop: 504.

FEATURES

source
 Location/Qualifiers
 1..504
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCGIM0242A10"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUCGIM library"
 /note="Vector: PMD42nv, Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson

ORIGIN

Query Match 72.6%; Score 19.6; DB 8; Length 504;
 Best Local Similarity 84.6%; Pred. No. 9.9e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 2 CGCTATGAGCAGCAGCGCCGAC 27
 Db 341 CGCATGAGCAGCAGCAGCGCAC 366

RESULT 13
 AM057282 599 bp mRNA linear EST 29-SEP-1999
 LOCUS ca03c06.y1 C elegans fem3 Q23 S1 Caenorhabditis elegans cDNA 5'
 DEFINITION similar to gb:y38f1a.a (ELEGANS);, mRNA sequence.
 ACCESSION AM057282
 VERSION AM057282.1 GI:5932921
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditioidea; Rhabditidae; Peloderae; Caenorhabditis.

REFERENCE
 AUTHORS 1 (bases 1 to 599)
 Ward, S., Smith, H., Clifton, S., Marra, M., Hillier, L., Kucaba, T.,
 Page, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Steptoe, M., Gibbons, M.,
 Harvey, N., Ritzer, E., Jackson, Y., McCann, R., Waterston, R. and
 Wilson, R.

TITLE UoFArizona-Washu C. elegans EST project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Samuel Ward, Ph.D.
 UoFArizona-Washu C. elegans EST project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watscn.wustl.edu
 Contact Harold Smith (hes@u.arizona.edu) for further information
 relating to organism, libraries, or clone availability.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 486.

FEATURES

source
 Location/Qualifiers
 1..599
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /db_xref="taxon:6239"
 /lab_host="DH5alpha cells"
 /clone_lib="C elegans fem3 Q23 S1"
 /note="Vector: pBluescript II SK+; Site 1: XhoI; Site 2:
 NotI; This C elegans library was made from fem-3(q23)-
 worms (produce only sperm at 25 C). cDNA was generated
 via oligo (dT) priming and directionally cloned into
 pBluescript II SK+ vector with a modified polylinker.
 This library was substracted with a second fem-1(hc17)
 (produce only oocytes at 25 C) library to enrich this

Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/dnares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PMD42 (GI:4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

library for sperm specific genes. First strand cDNA synthesis was primed with a NotI-15T oligo (sequence 5'-GACAGTTCGATGACGAGCGCGCCCTTTT-3'). After 2nd strand synthesis, XhoI-EcoRI adaptors (Stratagene #90120) were ligated, and cDNA was digested with NotI and ligated into XhoI-NotI digested vector."

ORIGIN

Query Match 72.6%; Score 19.6; DB 2; Length 599;
Best Local Similarity 84.6%; Pred. No. 9.9e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GCGCTCATGGAGCAGCAGCCCA 26
Db 494 GCGATCATGGAGCAGCAGCCCA 519

RESULT 14

LOCUS CB469087 617 bp mRNA linear EST 26-MAR-2003
DEFINITION CB469087 sm03.A04.f sn Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CB469087.1 GI:29275472
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 617)
Nellian, J.G., Kutish, G.F., Lu, Z., Zaak, A. and Rock, D.L.
Sequence analysis of African swine fever virus infected and non-infected porcine macrophage cDNA libraries
Unpublished (2003)
Contact: Nellian JG
Plum Island Animal Disease Center
US Department of Agriculture, Agricultural Research Service
PO Box 848, Greenport, NY 11944-848, USA
Tel: 631 323 3133
Fax: 631 323 3044
Email: jnellian@piadc.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim.alc option. Vector identified by cross match v0.990329 and Lucy v1.17p.
Seq primer: M13 Forward.

FEATURES

Location/Qualifiers
1..617
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="lymphoid"
/cell_type="macrophage"
/lab_host="DH10B"
/clone_id="sn"
/note="Vector: pSPORT1, site 1: NotI, site 2: SalI;
library made from pools of polyA selected RNA, isolated at different times post-infection (0 to 16 hrs) from African swine fever virus (isolate Pretoriuskop/96/4) infected swine macrophages. Macrophages were derived from peripheral blood mononuclear cells cultured for 48 hrs on plastic in the presence of 30% L929 supernatant."

ORIGIN

Query Match 72.6%; Score 19.6; DB 6; Length 617;
Best Local Similarity 84.6%; Pred. No. 9.9e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GCGTCATGGAGCAGCAGCCCA 27
Db 513 CACATCATGGAGCAGCAGCCCA 538

RESULT 15
A2999005/c

LOCUS A2999005 623 bp DNA linear GSS 27-APR-2001
DEFINITION 2M0266N12F Mouse 10kb plasmid UNGC2M library Mus musculus genomic
clone UNGC2M0266N12 F, genomic survey sequence.
ACCESSION A2999005
VERSION A2999005.1 GI:13870232
KEYWORDS
SOURCE
ORGANISM

REFERENCE

1 (bases 1 to 623)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A. and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

JOURNAL

Insert length: 10000 Std Error: 0.00
Plate: 0286 row: N column: 12
Seq primer: CGTTGTAAACGACGCGCAGCT
Class: plasmid ends
High quality sequence stop: 623.
Location/Qualifiers
1..623
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UNGCM0266N12"
/sex="Female"
/lab_host="B. coli strain XL10-Gold, T1-resistant, F-"
/clone_id="Mouse 10kb plasmid UNGC2M library"
/note="Vector: pMD42nv, Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (914732114|p|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptor complementary to the insert adaptor and purified. The sheared, adaptor mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

FEATURES

ORIGIN

Query Match 72.6%; Score 19.6; DB 8; Length 623;
Best Local Similarity 84.6%; Pred. No. 9.9e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GCGTCATGGAGCAGCAGCCCA 27
Db 460 GCGTCATGGAGCAGCAGCCCA 435

Search completed: December 15, 2004, 02:35:07
Job time : 89.0321 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:34:50 / Search time 84.2037 Seconds
(without alignments)
14040.270 Million cell updates/sec

Title: US-09-995-225B-42

Perfect score: 25

Sequence: 1 gaggcagtagtcacacactatg 25

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl:.*
1: gb_ba:.*
2: gb_htg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sy:.*
12: gb_un:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	AX498219	AX498219 Sequence
2	25	100.0	1343	AX254977	AX254977 Sequence
3	25	100.0	1536	AX664701	AX664701 Sequence
4	25	100.0	1826	AX463232	AX463232 Sequence
5	25	100.0	101882	AC021089	AC021089 Homo sapi
6	25	100.0	110000	AL139235	AL139235 Homo sapi
7	20.2	80.8	186606	AC129308	AC129308 Mus muscu
8	20	80.0	204798	AC125810	AC125810 Rattus no
9	20	80.0	236921	AC094151	AC094151 Rattus no
10	20	80.0	241131	AC118438	AC118438 Rattus no
11	20	80.0	245855	AC106203	AC106203 Rattus no
12	19.2	76.8	142525	AL671501	AL671501 Mouse DNA
13	19.2	76.8	149988	AL591602	AL591602 Human DNA
14	19.2	76.8	194457	AL831765	AL831765 Mouse DNA
15	19.2	76.8	219990	AC112943	AC112943 Mus muscu
16	19.2	76.8	249841	AC095767	AC095767 Rattus no
17	19.2	76.8	257767	AC095974	AC095974 Rattus no
18	19.2	76.8	326663	AF011889	AF011889 Human Xq2
19	19	76.0	1596	AF298642	AF298642 Avian pnc

C 20	19	76.0	1596	14	AF298643	AF298643 Avian pnc
C 21	19	76.0	1596	14	AF298644	AF298644 Avian pnc
C 22	19	76.0	1596	14	AF298645	AF298645 Avian pnc
C 23	19	76.0	1596	14	AF298646	AF298646 Avian pnc
C 24	19	76.0	1596	14	AF298647	AF298647 Avian pnc
C 25	19	76.0	1596	14	AF298648	AF298648 Avian pnc
C 26	19	76.0	1596	14	AF298649	AF298649 Avian pnc
C 27	19	76.0	1596	14	AF298650	AF298650 Avian pnc
C 28	18.8	75.2	11240	9	AC127902	AC127902 Homo sapi
C 29	18.8	75.2	50166	9	AL358053	AL358053 Human DNA
C 30	18.8	75.2	118632	9	AL133380	AL133380 Human DNA
C 31	18.8	75.2	156095	9	AC096577	AC096577 Homo sapi
C 32	18.8	75.2	170313	10	AC121973	AC121973 Mus muscu
C 33	18.8	75.2	176092	2	AC073252	AC073252 Homo sapi
C 34	18.8	75.2	181144	9	AL365444	AL365444 Human DNA
C 35	18.8	75.2	181717	9	AL353696	AL353696 Human DNA
C 36	18.8	75.2	209428	2	AC125948	AC125948 Rattus no
C 37	18.8	75.2	210321	2	AC132027	AC132027 Rattus no
C 38	18.6	74.4	86290	9	AC112503	AC112503 Homo sapi
C 39	18.6	74.4	110000	1	BX950851	Continuation (26 o
C 40	18.6	74.4	142488	10	AC087040	AC087040 Mus muscu
C 41	18.6	74.4	209139	2	AC083818	AC083818 Mus muscu
C 42	18.4	73.6	167123	10	AL663041	AL663041 Mouse DNA
C 43	18.4	73.6	208880	2	AC128103	AC128103 Rattus no
C 44	18.4	73.6	212178	2	AC125925	AC125925 Rattus no
C 45	18.4	73.6	220407	2	AC121626	AC121626 Rattus no

ALIGNMENTS

RESULT 1	AX498219	25 bp	DNA	linear	PAT 26-SEP-2002
LOCUS	AX498219	Sequence 42 from Patent WO0242461.			
DEFINITION	AX498219				
ACCESSION	AX498219.1	GI:23343141			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.				
TITLE	Endogenous and non-endogenous versions of human g protein-coupled				
JOURNAL	receptors				
FEATURES	-Patent: WO 0242461-A 42 30-MAY-2002;				
source	Arena Pharmaceuticals, Inc. (US)				
	Location/Qualifiers				
	1..25				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="Novel Sequence"				
ORIGIN					
Query Match	100.0%;	Score 25;	DB 6;	Length 25;	
Best Local Similarity	100.0%;	Pred. No. 0.25;			
Matches	25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1				
Db	1				
	1	GAGGCAGTAGTTCACACACTATG	25		
RESULT 2	AX254977	1343 bp	DNA	linear	PAT 10-OCT-2001
LOCUS	AX254977	Sequence 21 from Patent WO0170978.			
DEFINITION	AX254977				
ACCESSION	AX254977.1	GI:16074504			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
	Homo sapiens (human)				
	Homo sapiens				

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS 1 Taupier, R.J., Majumder, K., Spaderma, S.K., Smithson, G., Mezes, P.S. and Vernet, C.A.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0170978-A 21 27-SEP-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1. 1343
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCAGTAGTGTCCACACCTATGG 25
Db 1090 GAGGCAGTAGTGTCCACACCTATGG 1066
RESULT 3
AX664701/c AX664701 1526 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 10 from Patent WO02074960.
DEFINITION AX664701
ACCESSION AX664701 GI:29164461
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 Leiby, K.R., Kapeller-Libermann, R. and Glucksmann, M.
AUTHORS 38650, 26472, 5495, 65507, 81588 and 14354 methods and compositions of human proteins and uses thereof
TITLE Patent: WO 02074960-A 10 26-SEP-2002;
JOURNAL Millennium Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1. 1526
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
139. 1200
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD80046.1"
/db_xref="GI:29164462"
/translation="MEHTAHLAANSLWSFGSACGLGTVVTVYSLLQCLGLPAN
ILTVILISQVAROKSSVNYLLALAADLVLFVVDLFLEDFLNNQMPQVBDK
IIVLEESSHTSIWITVPLTIRYLAVCHPLKHTVSVPARTRKIVSVYITCFILTS
IPYWMNNTWTEDEYISFVSHVLIWVHCFTVGVPCSFILNSIVYKRRSGNRL
CGYSTGKTATLFTISFALIMAPRIMILVHAGAPIONRWLVIHMSDIAMMLALL
NTAINFLVYFISKRFRMTAAATLKAFKCKQKQPVQPTTHANFSITSSPWISFANSHC
IKMLVYQYDNGRPIKIVSP"
ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCAGTAGTGTCCACACCTATGG 25
Db 1227 GAGGCAGTAGTGTCCACACCTATGG 1203
RESULT 4
AX463232/c AX463232 1826 bp DNA linear PAT 15-JUN-2002
LOCUS Sequence 6 from Patent WO0248358.
DEFINITION

ACCESSION AX463232
VERSION AX463232.1 GI:21886202
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 Smolyar, A., Zhu, Z., Encinas, J., Watanabe, S. and Okigami, H.
AUTHORS Regulation of human chemokine-like receptor
TITLE Patent: WO 0248358-A 6 20-JUN-2002;
JOURNAL Bayer Aktiengesellschaft (DE)
FEATURES Location/Qualifiers
source 1. 1826
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 1826;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCAGTAGTGTCCACACCTATGG 25
Db 1132 GAGGCAGTAGTGTCCACACCTATGG 1108
RESULT 5
AC021089 AC021089 101882 bp DNA linear PRI 22-MAR-2003
LOCUS Homo sapiens chromosome 16 clone CTD-2264D9, complete sequence.
DEFINITION AC021089
ACCESSION AC021089 GI:29150340
VERSION
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
5 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Mar 22, 2003 this sequence version replaced gi:20330797.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory
www-shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.

FEATURES
source
1. 101882
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTD-2264D9"

ORIGIN
Query Match 100.0%; Score 25; DB 9; Length 101882;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCGACTGTTCCACACTATGC 25
|||||
Db 18207 GAGCGACTGTTCCACACTATGC 18231

RESULT 6
AL139235.0
WPCOMMENT
Sequence split into 4 fragments LOCUS AL139235 Accession AL139235
Fragment Name Begin End
AL139235.0 1 11000
AL139235-1 100001 210000
AL139235-2 200001 310000
AL139235-3 300001 407025
AL139235 407025 bp DNA 1linear HTG 16-JUL-2001
DEFINITION Homo sapiens chromosome 9 clone RP11-3215, 31 unordered pieces.
ACCESSION AL139235
VERSION AL139235.22 GI:11034497
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Plumb, B.
Direct Submission
Submitted (14-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requesters: clonerequest@sanger.ac.uk
On Oct 26, 2000 this sequence version replaced gi:11024977.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: DA3215
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 398994 bases at least Q40
Consensus quality: 395933 bases at least Q30
Consensus quality: 400050 bases at least Q20
Insert size: 174695; 17.7% error; agarose-fp
Quality coverage: 4.93x in Q20 bases; sum-of-contigs Quality
coverage: 14.49x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2425: contig of 2425 bp in length
* 2426 2525: gap of 100 bp
* 2526 10023: contig of 7498 bp in length

10024 10123: gap of 100 bp
10124 36364: contig of 26241 bp in length
36365 43873: gap of 100 bp
43874 43973: contig of 7409 bp in length
43974 49988: contig of 6015 bp in length
49989 50088: gap of 100 bp
50089 93603: contig of 43515 bp in length
93604 93703: gap of 100 bp
93704 96191: contig of 2488 bp in length
96192 96291: gap of 100 bp
96292 99433: contig of 3142 bp in length
99434 99533: gap of 100 bp
99534 101556: contig of 2023 bp in length
101557 101657: gap of 100 bp
101657 184109: contig of 82453 bp in length
184110 184209: gap of 100 bp
184210 290696: contig of 106487 bp in length
290697 290796: gap of 100 bp
290797 292892: contig of 2096 bp in length
292893 292992: gap of 100 bp
292993 296019: contig of 3026 bp in length
296019 296119: gap of 100 bp
296119 322154: contig of 26036 bp in length
322155 322255: gap of 100 bp
322255 324646: contig of 2392 bp in length
324647 324746: gap of 100 bp
324747 329069: contig of 4323 bp in length
329070 329169: gap of 100 bp
329170 331371: contig of 2202 bp in length
331372 331471: gap of 100 bp
331472 333909: contig of 2438 bp in length
333910 334009: gap of 100 bp
334010 369445: contig of 35436 bp in length
369446 369545: gap of 100 bp
369546 371996: contig of 2451 bp in length
371997 372096: gap of 100 bp
372097 375925: contig of 3829 bp in length
375926 376025: gap of 100 bp
376026 378132: contig of 2107 bp in length
378133 378232: gap of 100 bp
378233 381590: contig of 3358 bp in length
381591 381690: gap of 100 bp
381691 383726: contig of 2036 bp in length
383727 383826: gap of 100 bp
383827 390832: contig of 7006 bp in length
390833 390932: gap of 100 bp
390933 394216: contig of 3284 bp in length
394217 394316: gap of 100 bp
394317 396358: contig of 2042 bp in length
396359 396458: gap of 100 bp
396459 398733: contig of 2275 bp in length
398734 398833: gap of 100 bp
398834 402385: contig of 3552 bp in length
402386 402485: gap of 100 bp
402486 404522: contig of 2037 bp in length
404523 404622: gap of 100 bp
407025: contig of 2403 bp in length.
Location/Qualifiers
1. 407025
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-3215"
/clone_11b="RP11-3215"
1. 2425
/note="assembly fragment: 05385
fragment_chain: 1"
2526. 10023
/note="assembly fragment: 03131
fragment_chain: 1"
10124. 36364
misc_feature
misc_feature
misc_feature

```

misc_feature /note="assembly_fragment:06402
fragment_chain:1"
3645. .43873
/note="assembly_fragment:01436
fragment_chain:1"
43974. .43988
/note="assembly_fragment:04547
fragment_chain:1"
50089. .53603
/note="assembly_fragment:01636
fragment_chain:1"
93704. .96191
/note="assembly_fragment:02957
fragment_chain:2"
96292. .99433
/note="assembly_fragment:00499
fragment_chain:2"
99534. .101556
/note="assembly_fragment:03112
fragment_chain:2"
101657. .184109
/note="assembly_fragment:01419
fragment_chain:3"
184210. .290696
/note="assembly_fragment:00925
fragment_chain:3"
290797. .292892
/note="assembly_fragment:03065
fragment_chain:4"
292993. .296018
/note="assembly_fragment:05057
fragment_chain:4"
296119. .322154
/note="assembly_fragment:00269"
322255. .324646
/note="assembly_fragment:00666"
324747. .329069
/note="assembly_fragment:01113"
329170. .331371
/note="assembly_fragment:01461"
331472. .333909
/note="assembly_fragment:01516"
334010. .369445
/note="assembly_fragment:01942.0"
369546. .371996
/note="assembly_fragment:02194"
372097. .375925
/note="assembly_fragment:02249"
376026. .378132
/note="assembly_fragment:02485"
378233. .381590
/note="assembly_fragment:02649"
381691. .383726
/note="assembly_fragment:02869"
383827. .390832
/note="assembly_fragment:03540"
390933. .394216
/note="assembly_fragment:04405"
394317. .396358
/note="assembly_fragment:04595"
396459. .398733
/note="assembly_fragment:05768"
398834. .402385
/note="assembly_fragment:06106"
402486. .404522
/note="assembly_fragment:06204"
404623. .407025
/note="assembly_fragment:06292"

```

ORIGIN

```

Query Match 100.0%; Score 25; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 GAGGACGATGTTCCACACCTATGG 25
Db 34226 GAGGACGATGTTCCACACCTATGG 34250

RESULT 7
AC129308
LOCUS
DEFINITION
Mus musculus BAC clone RP24-308E2 from chromosome 14, complete
sequence.
AC129308
VERSION
AC129308.6 GI:33285230
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 186606)
Wang, C., Abbot, A. and Haglund, K.
The sequence of Mus musculus BAC clone RP24-308E2
2 (bases 1 to 186606)
Wilson, R.
Sequencing of Mus musculus
3 (bases 1 to 186606)
Unpublished (2001)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (28-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 186606)
Wilson, R.K.
Direct Submission
Submitted (11-JUN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 186606)
Wilson, R.K.
Direct Submission
Submitted (26-JUL-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 186606)
Wilson, R.
Direct Submission
Submitted (27-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 26, 2003 this sequence version replaced gi:31581805.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.wustl.edu
----- Summary Statistics
Center project name: M_BB0308E02
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see

http://genome.wustl.edu

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (http://www.chori.org) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES
source Location/Qualifiers

```
1..186606
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="takara:10090"
/chromosome="14"
/map="14"
/clone="RP24-308B2"
/clone_1b="RPCI-24"
221..574
/rpc_family="MaLR"
repeat_region
/rpc_family="MaLR"
600..715
/rpc_family="Alu"
repeat_region
1550..1674
/rpc_family="B4"
2438..3434
/rpc_family="L1"
3575..3932
/rpc_family="MaLR"
4945..5054
/rpc_family="MaLR"
misc_feature
5884..7189
/note="Sequence derived from PCR product of project DNA."
repeat_region
/rpc_family="L1"
6134..6199
/note="Sequence derived from one plasmid subclone."
6268..6377
/note="Sequence derived from one plasmid subclone."
7705..8093
/rpc_family="ERVK"
repeat_region
8952..9065
/rpc_family="ERV1"
repeat_region
9997..10194
/rpc_family="B4"
10232..10309
/rpc_family="ID"
repeat_region
12635..12695
/rpc_family="ID"
repeat_region
13111..13466
/rpc_family="MaLR"
14142..14561
/rpc_family="MaLR"
repeat_region
14566..14764
/rpc_family="MaLR"
repeat_region
15512..15615
/rpc_family="B4"
15616..15770
/rpc_family="MaLR"
repeat_region
15772..16030
/rpc_family="L1"
repeat_region
16031..16100
/rpc_family="B2"
16073..16177
/rpc_family="B2"
repeat_region
16845..17117
/rpc_family="L1"
repeat_region
25187..25478
/rpc_family="MaLR"
25573..25643
/rpc_family="MaLR"
repeat_region
25929..26013
/rpc_family="L1"
26476..26521
/rpc_family="B4"
repeat_region
```

```
repeat_region 26549..26672
/rpc_family="B4"
repeat_region 26905..27276
/rpc_family="MaLR"
repeat_region 27341..27421
/rpc_family="B4"
repeat_region 28427..28713
/rpc_family="MaLR"
repeat_region 29106..29587
/rpc_family="ERV1"
repeat_region 29608..30271
/rpc_family="L1"
repeat_region 32786..33794
/rpc_family="L1"
repeat_region 33950..35375
/rpc_family="L1"
repeat_region 35376..36321
/rpc_family="RMR1A"
repeat_region 36322..38311
/rpc_family="L1"
repeat_region 38372..39557
/rpc_family="L1"
unaure 39558..39870
/rpc_family="L1"
/note="Unresolved simple sequence repeat."
repeat_region 40130..40209
/rpc_family="Alu"
repeat_region 40194..40421
/rpc_family="MaLR"
repeat_region 41371..41425
/rpc_family="ERV1"
repeat_region 42893..43083
/rpc_family="B2"
repeat_region 46442..46597
/rpc_family="B2"
repeat_region 46738..47109
/rpc_family="B2"
repeat_region 47267..47414
/rpc_family="MaLR"
repeat_region 48365..48497
/rpc_family="Alu"
repeat_region 49562..49772
/rpc_family="B4"
repeat_region 51468..51582
/rpc_family="B2"
repeat_region 54385..54626
/rpc_family="Alu"
repeat_region 55038..55103
/rpc_family="RMR15"
repeat_region 55595..55743
/rpc_family="ERV1"
repeat_region 55909..55998
/rpc_family="Achobo"
repeat_region 56001..56182
/rpc_family="Alu"
repeat_region 57800..57881
/rpc_family="MER1_type"
repeat_region 59721..59939
/rpc_family="MaLR"
repeat_region 60010..60335
/rpc_family="B2"
repeat_region 61776..61870
/rpc_family="Ricksna"
/rpc_family="MaLR"
repeat_region
```

Query Match 80.8%; Score 20.2; DB 10; Length 186606;
Best Local Similarity 88.0%; Pred.No. 32;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAGGCACTACTGTCACCACTATGCG 25
DB 11111 GAGGCACTAGTGTACACCACTATGCG 11135

RESULT 8

AC125810/c	204798 bp	DNA	linear	HTG 20-NOV-2002	
LOCUS					
DEFINITION	Rattus norvegicus clone CH230-345P21, WORKING DRAFT SEQUENCE.				
ACCESSION	AC125810				
VERSION	AC125810.3 GI:25139743				
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 204798)				
AUTHORS	Muzny, D., Marie, M., Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, S., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhammed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Ditya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Bayes, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, U., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kover, C., Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshuber, L., Loulsegod, H., Lozado, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaekelam, O., Okwona, G., Olarnpusagoon, A., Pal, S., Parks, K., Paeternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poldexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., A., Reigh, R., Rell, B., Rell, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Slason, I., Sitter, C., D., Smajs, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, U., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D., R., Holt, R., A., Smith, H., O., Weinstock, G., and Gibbs, R., A.				
TITLE	Direct Submission				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 204798)				
AUTHORS	Worley, K., C.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
REFERENCE	3 (bases 1 to 204798)				
AUTHORS	Rat Genome Sequencing Consortium.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-NOV-2002) Human Genome Sequencing Center, Department				

COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Nov 20, 2002 this sequence version replaced gi:23915433.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KADO

Center clone name: CH230-345P21

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 190846 bases at least Q40

Consensus quality: 192519 bases at least Q30

Consensus quality: 193364 bases at least Q20

Estimated insert size: 194659; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 204798: contig of 204798 bp in length.

Location/Qualifiers

1..204798

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-345P21"

886..1703

/note="clone boundary"

clone_end:77

site:

end_sequence: BZ131587"

complement(196726..197574)

/note="clone boundary"

clone_end:Sp6

site:

end_sequence: BZ131588"

203715..204798

/note="wgs end_extension"

clone_end:Sp6"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

RESULT 10
AC118438
LOCUS
DEFINITION
AC118438 241131 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-236A23, *** SEQUENCING IN PROGRESS

AC118438
AC118438.7 GI:25009656
HTG, HTGS_PHASE2, HTGS_DRAFT, HTGS_ENRICHED.
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 241131)
Mizny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angulano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, U., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garner, P.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gumarate, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, U.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
Hollins, B., Howell, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., Khan, J., Kovar, C.,
Kowitz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshuber, L., Louised, H., Lozano, R.U., Lu, X., Ma, U.,
Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, B.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackelmehe, O., Okunolu, G., Olarinmoye, A., Pal, S., Parks, K.,
Pasernak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L.-L.,
Punzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Riz, S.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajb, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabors, Z., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Triolo, P., Uman, K.,
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Zhao, S., Dunn, D., von
Weinhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 241131)
Worley, K.C.
REFERENCE
Submitted (17-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
3 (bases 1 to 241131)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23810762.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTUD
Center clone name: CH230-236A23
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 229030 bases at least Q40
Consensus quality: 231946 bases at least Q30
Consensus quality: 23305 bases at least Q20
Estimated insert size: 234677; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 241131: contig of 241131 bp in length.
Location/Qualifiers
1..241131
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-236A23"
1..1720
/note="wgs end extension
clone_end:5p6"
complement(3965..4865)
/note="clone_boundary
clone_end:5p6
site:
end_sequence:BN096525"
156555..218103
/note="clone_boundary
clone_end:T7
site:
end_sequence:BN096524"
223741..224741
/note="wgs end extension
clone_end:T7"

ORIGIN
Query Match 80.0%; score 20; DB 2; length 241131;
Best Local Similarity 100.0%; Pred. No. 39;

Matches 20, Conservative 0, Mismatches 0, Indels 0, Gaps 0

Qy 2 AGGAGTAGTGGACACT 21
 |||||
 Db 202325 AGGAGTAGTGGACACT 202344

RESULT 11
 AC106203
 LOCUS
 DEFINITION
 AC106203 245855 bp DNA linear HTG 13-MAY-2003
 Rattus norvegicus clone CH230-27C8, *** SEQUENCING IN PROGRESS ***
 8 unordered pieces.
 AC106203
 AC106203.4 GI:30579094
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS
 1 (bases 1 to 245855)
 Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bismail, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Faller, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Gavara, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, A., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Haviak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huily, S., Hume, J., Idlebird, D., Jackson, A., Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowik, C., Kraft, C.L., Lepow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzuela, L., Loulsegod, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okunolu, G., Olarnpunsegon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richard, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shaatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smales, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, B., Umanil, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, J., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Woodan, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von Niederhuesen, A., Weis, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE
 JOURNAL
 REFERENCE
 2 (bases 1 to 245855)

AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Worley, K.C.
 Direct Submission
 Submitted (12-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 245855)
 Rat Genome Sequencing Consortium.
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 this sequence version replaced gi:22269037.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GJVB
 Center clone name: CH230-27C8
 ----- Summary Statistics
 Assembly program: Atlas 3.0.
 Consensus quality: 212523 bases at least Q40
 Consensus quality: 217649 bases at least Q30
 Consensus quality: 221314 bases at least Q20
 Estimated insert size: 22736; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 32679: contig of 32679 bp in length
 32680 32779: gap of unknown length
 32780 54985: contig of 22206 bp in length
 54986 55085: gap of unknown length
 55086 73449: contig of 18364 bp in length
 73450 73549: gap of unknown length
 73550 211048: contig of 137499 bp in length
 211049 211148: gap of unknown length
 211149 221087: contig of 9939 bp in length
 221088 221187: gap of unknown length
 221188 239949: contig of 18762 bp in length
 239950 240049: gap of unknown length
 240050 241182: contig of 1133 bp in length
 241183 241282: gap of unknown length
 241283 245855: contig of 4573 bp in length.

Location/Qualifiers
 1. 245855
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-27C8"
 73550..74985
 /note="wgs_contig"

FEATURES
 source
 misc_feature

misc_feature 75386..77137
/note="wgs_config"
misc_feature 22188..222913
/note="wgs_contig"

ORIGIN

Query Match 80.0%; Score 20; DB 2; Length 24585;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGGCACTAGTGGCACACCT 21
|||||
|||||

Db 138907 AGGCACTAGTGGCACACCT 138926

RESULT 12
AL671501 142525 bp DNA 11linear ROD 01-JUL-2003
LOCUS Mouse DNA sequence from clone RP23-313P3 on chromosome X, complete
DEFINITION
ACCESSION AL671501
VERSION AL671501 GI:32400078
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 142525)
Whitehead, S.
Direct Submission
Submitted (01-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Jul 1, 2003 this sequence version replaced gi:32165975.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
EM: EMBL; SW: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPI-23 Mouse BAC library
constructed by the group of Pletier de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBAC3.6.
Location/Qualifiers
1. 142525
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"

FEATURES
source

/clone="RP23-313P3"
/clone_lib="RP23-23"

ORIGIN

Query Match 76.8%; Score 19.2; DB 10; Length 142525;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGGCACTAGTGGCACACCTATGG 25
|||||
|||||

Db 103084 AGGCACTAGTGGCACACCTATGG 103107

RESULT 13
AL591602 149988 bp DNA 11linear PRI 16-DEC-2001
LOCUS Human DNA sequence from clone RP5-1092E23 on chromosome 1, complete
DEFINITION
ACCESSION AL591602
VERSION AL591602.7 GI:17907262
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 149988)
Wallis, J.
Direct Submission
Submitted (16-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Dec 18, 2001 this sequence version replaced gi:16973109.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: EM: EMBL; SW:
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone configs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
RP5-1092E23 is from the library RPI-5 constructed by the group of
Pletier de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RP5-1092E23 it may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP5-1092E23 is at 149988 in this
sequence. The true right end of clone RP4-573J3 is at 2000 in this
sequence.
Location/Qualifiers
1. 149988
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP5-1092E23"
/clone_lib="RPI-5"

FEATURES
source

ORIGIN

Query Match 76.8%; Score 19.2; DB 9; Length 149988;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AGGCAGTAGTTGCCACACTATG 25
|||||
15256 AGGCAGTAGTTGCCACCTCTAG 15279

RESULT 14

AL831765 194457 bp DNA linear ROD 24-SEP-2002
LOCUS Mouse DNA sequence from clone RP23-148012 on chromosome 4, complete
DEFINITION
sequence.

ACCESSION AL831765
VERSION AL831765.3 GI:23320939
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 194457)

TITLE Direct Submission
JOURNAL Submitted (24-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Sep 25, 2002 this sequence version replaced gi:22798457.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL, Sw,
SWISSPROT; Tr, TrEMBL; Wp, WormPeP; Information on the WormPeP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-148012 is
from the RPI-23 Mouse PAC library
constructed by the group of Pieter de Jong
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6.

FEATURES

source

1. 194457
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-148012"
/clone_1lb="RPI-23"

ORIGIN

Query Match 76.8%; Score 19.2; DB 10; Length 194457;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AGGCAGTAGTTGCCACACTATG 25
|||||
|||||

Db 155036 AGGCATAGTTCCACACTATG 155059

RESULT 15

AC112943/c 219990 bp DNA linear ROD 21-OCT-2003
LOCUS Mus musculus chromosome 16, clone RP23-424C7, complete sequence.
DEFINITION
AC112943
AC112943.12 GI:37777390
KEYWORDS HTG.

SOURCE

Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 219990)
2 (bases 1 to 219990)

REFERENCE
AUTHORS Birren, B., Nussbaum, C. and Lander, E.
JOURNAL Unpublished
TITLE Mus musculus chromosome 16, clone RP23-424C7
AUTHORS 2 (bases 1 to 219990)

Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Baatien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagot, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lahocque, K., Lamazares, R.,
Lander, T., Lehocsky, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M.,
McGowan, P., McKernan, K., Meldrum, J., Menes, L., Mihova, T.,
Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollard, V., Roman, J.,
Rella, R., Rieback, M., Riley, R., Rhee, C., Rogov, P., Roman, J.,
Roetli, M., Roy, A., Santos, R., Schauer, S., Schuppach, R., Seaman, S.,
Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talama, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travia, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 219990)

REFERENCE
AUTHORS Birren, B., Nussbaum, C. and Lander, E.
JOURNAL Direct Submission
TITLE Submitted (04-SEP-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 219990)

Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arachchi, H.M., Barna, N., Baatien, V., Bloom, T.,
Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearrellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hatz, N., Hagopian, D., Hago, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lander, E., Levine, R.,
Lindblad-Toh, K., Liu, G., Liu, X., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menes, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Rella, R., Rhee, C., Rogov, P.,
Roman, J., Schauer, S., Schuppach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talama, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

COMMENT

On Oct 21, 2003 this sequence version replaced gi:34447307.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L20313

Center clone name: 424_C_7

FEATURES

source

----- Location/Qualifiers

1..219990

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="16"

/map="16"

/clone="RP23-424C7"

/clone_id="RP23-424C7" Female Mouse BAC"

misc_feature

1..15627

/note="wgs_end_extension"

clone_end:SP6"

1..1056

/rpt_family="Lx6"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

10727..10875

/rpt_family="B1_MM"

repeat_region

11642..11693

/rpt_family="AT_rich"

repeat_region

11992..12028

/rpt_family="CNGA)n"

repeat_region

complement(13169..15310)

misc_feature

15028..15033

/note="clone boundary"

clone_end:SP6

site:EcoRI"

15675..15912

/rpt_family="CATATA)n"

complement(15915..16226)

/rpt_family="MTC"

complement(16878..16993)

/rpt_family="MTC"

17006..18011

/rpt_family="ORR1B-int"

18012..18383

/rpt_family="RLTR10"

18391..18678

/rpt_family="ORR1B-int"

18679..18730

/rpt_family="ORR1D"

18766..18942

/rpt_family="CATATA)n"

18992..19178

/rpt_family="CATATA)n"

19227..19408

/rpt_family="RLTR1"

19426..19743

/rpt_family="RLTR1"

complement(19801..20141)

/rpt_family="MT2B"

20411..20809

/rpt_family="RLTR9C"

complement(20817..23539)

/rpt_family="MERVA"

complement(23544..23788)

/rpt_family="MT2A"

23795..24137

/rpt_family="ORR1C"

complement(24138..24444)

/rpt_family="L1"

24445..24486

/rpt_family="L1"

complement(24487..24525)

/rpt_family="TG)n"

complement(24525..24895)

/rpt_family="L1"

24895..25932

/rpt_family="Lx2B"

25978..30421

/rpt_family="L1_MM"

complement(30422..30569)

/rpt_family="B1_MM"

30570..31592

/rpt_family="L1_MM"

complement(31593..31918)

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

Query Match

Beat Local Similarity 76.8%; Score 19.2; DB 10; Length 219990;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 2 AGGAGTAGTGGCCACCTATGG 25

DB 198507 AGGAGTAGTGGCCACCTTAG 198484

Search completed: December 15, 2004, 00:42:58
Job time : 88.2037 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:00:04 ; Search time 11.7084 Seconds
(without alignments)
11208.653 Million cell updates/sec

Title: US-09-995-225B-42

Perfect score: 25

Sequence: 1 gagcgagtagtcgcacactatg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001s:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	6	ABT04897
2	25	100.0	1202	6	ABV73364
3	25	100.0	1343	4	AA515731
4	25	100.0	1343	10	ADJ87768
5	25	100.0	1343	12	ADJ79324
6	25	100.0	1343	12	ADJ79324
7	25	100.0	1526	6	AA046858
8	25	100.0	1526	8	ACA60889
9	25	100.0	1526	10	ABN84272
10	25	100.0	1826	6	ABN84272
11	25	100.0	2273	12	ADO28955
12	18.2	72.8	295	4	AAH57320
13	18.2	72.8	147309	6	ABK49450
14	17.6	70.4	879	4	AA194220
15	17.6	70.4	1376	12	ADO60492
16	17.6	70.4	152501	12	ADP67269
17	17.4	69.6	208	5	ADJ73980
18	17.4	69.6	208	5	ADJ73980
19	17.4	69.6	248	5	ADJ67597
20	17.4	69.6	350	3	AA08203
21	17.4	69.6	527	5	ADJ39217

C	22	17.4	69.6	674	4	AA531458
C	23	17.4	69.6	674	6	AB06782
C	24	17.4	69.6	674	10	AD010804
C	25	17.4	69.6	1182	6	AB222351
C	26	17.4	69.6	1596	10	ADP68972
C	27	17.4	69.6	1596	10	ADP89293
C	28	17.4	69.6	1596	12	ADJ96792
C	29	17.4	69.6	1596	12	ADJ07739
C	30	17.4	69.6	1596	12	ADP67263
C	31	17.4	69.6	1614	3	AA14952
C	32	17.4	69.6	1644	12	ADQ30909
C	33	17.4	69.6	1666	10	ADP68973
C	34	17.4	69.6	1666	10	ADP89294
C	35	17.4	69.6	1666	12	ADJ96793
C	36	17.4	69.6	1666	12	ADJ07740
C	37	17.4	69.6	1666	12	ADP67264
C	38	17.4	69.6	1860	10	ADP68975
C	39	17.4	69.6	1860	10	ADP89296
C	40	17.4	69.6	1860	12	ADJ96795
C	41	17.4	69.6	1860	12	ADJ07742
C	42	17.4	69.6	1860	12	ADP67266
C	43	17.4	69.6	2162	4	AA531353
C	44	17.4	69.6	2162	6	AB066677
C	45	17.4	69.6	2162	10	AD010699

ALIGNMENTS

RESULT 1	ABT04897	standard; DNA; 25 BP.
ID	ABT04897	
AC	ABT04897	
XX	DT	11-OCT-2002 (first entry)
XX	DE	Human G protein coupled receptor hrup35 PCR primer SEQ ID NO: 42.
XX	KW	Human; G-protein coupled receptor; GPCR; hrup28; hrup29; hrup30; hrup31;
XX	KW	hrup32; hrup33; hrup34; hrup35; hrup36; hrup37; PCR; primer; ss.
XX	OS	Homo sapiens.
XX	PN	W0200242461.A2.
XX	PD	30-MAY-2002.
XX	PF	26-NOV-2001; 2001WO-US044386.
XX	PR	27-NOV-2000; 2000US-0253404P.
XX	PR	12-DEC-2000; 2000US-0253466P.
XX	PR	20-FEB-2001; 2001US-0270266P.
XX	PR	20-FEB-2001; 2001US-0270286P.
XX	PR	06-APR-2001; 2001US-0282032P.
XX	PR	06-APR-2001; 2001US-0282356P.
XX	PR	06-APR-2001; 2001US-0282358P.
XX	PR	06-APR-2001; 2001US-0282359P.
XX	PR	14-MAY-2001; 2001US-0290917P.
XX	PR	31-JUL-2001; 2001US-0309208P.
XX	PA	(AREN-) ARENA PHARM INC.
XX	PI	Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX	DR	WPI; 2002-566565/60.
XX	PT	Novel endogenous and non-endogenous versions of G protein-coupled
XX	PT	receptor useful for identification of candidate compounds as receptor
XX	PT	agonists or antagonists for use as therapeutic agents.
XX	PS	Example 1; Page 28; 84pp; English.

CC The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a PCR primer used to isolate a GPCR coding sequence of the invention
XX
SQ Sequence 25 BP; 6 A; 6 C; 8 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTGGCCACACCTATGG 25
Db 1 GAGGCACTAGTGGCCACACCTATGG 25

RESULT 2
ABV73364/C
ID ABV73364 standard; DNA; 1202 BP.

XX ABV73364;
AC
XX 22-JAN-2003 (first entry)
DT
XX Human TGR20 polypeptide encoding DNA.
DE
XX G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory;
KW antileptic; nootropic; neuroprotective; antianemic; antitumor; human;
KW antiparkinsonian; antileptic; TGR20; gene; de.
XX
XX Homo sapiens.

OS
XX Key Location/Qualifiers
FH 68..1129
FT CDS /tag= a
FT /product= "TGR20"
FT /note= "GPCR polypeptide"
XX
XX WO200277001-A2.

XX 03-OCT-2002.
XX 08-MAR-2002; 2002WO-US007171.
XX 09-MAR-2001; 2001US-00802803.
XX 16-MAR-2001; 2001US-0276649P.
XX
XX (TULA-) TULARIK INC.

XX Tian H, Zhao J, Chen J, Cutler G;
PI
XX WPI; 2003-018881/01.
DR P-PSDB; ABB82499.

XX New G-protein coupled receptor polypeptides and polynucleotides useful
PT for identifying compounds for treating a TGR-associated disorder, e.g.
PT psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's
PT disease, anemia.
XX
XX Claim 6; Page 63; 87pp; English.

XX The invention relates to G-protein coupled receptor (GPCR) polypeptides
CC and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183,
CC TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are
CC useful for identifying compounds for treating a TGR-associated disorder,
CC such as psoriasis, inflammatory bowel disease, hyperlipidemia,
CC Parkinson's disease, Huntington's disease, anemia, immune and blood
CC disorders, ulcerative colitis, Crohn's disease or spleen enlargement.
CC They are also useful for identifying cells such as kidney, liver,
CC hypothalamus, colon, adipose, or spleen cells, for forensics and
CC paternity determination, diagnosing diseases and examining signal
CC transduction. The present sequence represents a human TGR20 polypeptide

CC encoding DNA
XX
SQ Sequence 1202 BP; 267 A; 380 C; 258 G; 297 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 8; Length 1202;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTGGCCACACCTATGG 25
Db 1156 GAGGCACTAGTGGCCACACCTATGG 1132

RESULT 3
AAS15731/C
ID AAS15731 standard; cDNA; 1343 BP.

XX AAS15731;
AC
XX 14-FEB-2002 (first entry)
DT
XX DNA encoding chemokine receptor family related protein, NOV11.

DE
XX NOV; cytosolic; psoriasis; nootropic; neuroprotectant;
KW cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;
KW haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;
KW differentiation; proliferation; haematopoiesis; wound healing;
KW angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;
KW haemophilia; allergy; pendred syndrome; skeletal dysplasia;
KW ischaemic injury; neuroepithelial disorder; hepatitis; heart failure;
KW chemokine receptor; chromosome 1; ss.
XX
XX Homo sapiens.

OS
XX Key Location/Qualifiers
FH 1
FT 5'UTR /tag= a
FT 2..1063
FT CDS /tag= b
FT /product= "Chemokine receptor related protein, NOV11"
FT sig_peptide 2..142 C
FT /tag= c
FT mat_peptide 143..1060
FT /tag= d
FT /note= "Mature chemokine receptor related protein, NOV11"
FT 3'UTR /tag= e
XX
XX WO200170978-A2.

XX 27-SEP-2001.
XX 20-MAR-2001; 2001WO-US009093.

XX 20-MAR-2000; 2000US-0190768P.
XX 20-MAR-2000; 2000US-0190835P.
XX 22-MAR-2000; 2000US-0190972P.
XX 22-MAR-2000; 2000US-0191199P.
XX 24-MAR-2000; 2000US-0191947P.
XX 28-MAR-2000; 2000US-0192657P.
XX 28-MAR-2000; 2000US-0192664P.
XX 28-MAR-2000; 2000US-0192665P.
XX 28-MAR-2000; 2000US-0192844P.
XX 29-MAR-2000; 2000US-0192836P.
XX 31-MAR-2000; 2000US-0193843P.

XX (CURA-) CURAGEN CORP.

XX Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;
PI Vernet CM;
XX WPI; 2001-639127/73.
DR P-PSDB; AAU10068.

XX polypeptides and nucleic acids related to chloride channel, insulin-like
PT growth factor family of proteins, useful for diagnosing and treating
PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
XX
XX Claim 9; Page 44; 151pp; English.
XX
XX The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX
CC polypeptides are useful for treating pathology associated with NOVX
CC polypeptide, determining the presence of or predisposition to a disease
CC associated with altered levels of NOVX, identifying agents binding to
CC NOVX and treatment of disorders associated with altered expression of
CC members of chloride channel-associated proteins e.g. cystic fibrosis and
CC congenital myotonia. NOVX proteins are useful in treatment of disorders
CC including porphyria, cancer, diabetes, metabolic disorders of pancreas,
CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
CC in one or more organs (e.g. haemophilia, anaemia), Pendred syndrome,
CC skeletal dysplasias, disorders characterised by altered cell shape,
CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial
CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
CC treatment of disorders of vascular smooth muscle cell differentiation,
CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are
CC useful to screen for molecules which inhibit or enhance NOVX activity or
CC function and are useful as targets for the identifying small molecules,
CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV
CC sequences are also useful for: identifying a cell or tissue type in a
CC biological sample; amplifying DNA sequences from very small biological
CC samples e.g. hair or skin or body fluids and as primers and probes to
CC identify and/or clone NOVX homologues. NOVX proteins are useful
CC immunogens to generate antibodies to monitor protein levels and modulate
CC NOVX activity. Cells comprising the nucleic acids are useful for
CC producing transgenic animals, for studying the function and/or activity
CC of NOVX protein and identifying and/or evaluating modulators of NOVX
CC protein activity. This sequence encodes NOV11 (located on chromosome 1)
CC related to the chemokine receptor family of proteins, one of 12 NOV
CC polypeptides described in the method of the invention
XX
XX Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 25; DB 4; Length 1343;
Beat Local Similarity 100.0%; Pred. No. 0.048;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCACTAGTGGCACACCTATGG 25
DB 1090 GAGGCACTAGTGGCACACCTATGG 1066

RESULT 4
ADJ87768/c
ID ADJ87768 atandard; DNA; 1343 BP.
XX
XX ADJ87768;
XX
XX 06-MAY-2004 (first entry)
XX
XX G-coupled protein receptor-related protein coding sequence #57.
XX
XX novel protein; G-coupled protein receptor-related protein;
KM cardiomyopathy; atherosclerosis; cell signal processing-related disorder;
KM metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
KM Huntington's disease; epilepsy; pain; hypercholesterolaemia;
KM obesity; hypertension; Crohn's disease; systemic lupus erythematosus;
KM viral infections; bacterial infection; parasitic infection;
KM hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;
KM Alzheimer's disease; tuberculous atherosclerosis; hypercalcaemia; cerebral palsy;
KM gene; ds.
XX
XX Unidentified.
XX
XX OS
XX PN
XX MO2002102321-A2.

PD 27-DEC-2002.
XX
XX 18-JUN-2002; 2002MO-US019522.
PF
XX 18-JUN-2001; 2001US-0298994P.
PR 18-JUN-2001; 2001US-0299134P.
PR 04-OCT-2001; 2001US-0097244G.
PR 06-JUN-2002; 2002US-00299134.
PR 07-JUN-2002; 2002US-00298994.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;
PI Pena CE, Spaderna SK, Zhong M;
XX WPI; 2003-167441/16.
DR P-PSDB; ADJ87769.
XX
XX New MOLX polypeptides and polynucleotides, useful in gene therapy,
PT particularly for treating or preventing e.g. cardiomyopathy,
PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
PT cancer, stroke or pain.
XX
XX Claim 8; SEQ ID NO 203; 378pp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel G-
CC coupled protein receptor-related (MOL) proteins. The DNA and protein
CC sequences of the invention are useful for treating or preventing a MOL-
CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
CC associated with cell signal processing and metabolic pathway modulation,
CC or diabetes. The DNA and protein sequences are also useful for the
CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
CC pain, hypercholesterolaemia, obesity, hypertension, Crohn's disease,
CC systemic lupus erythematosus, viral infections, bacterial infections,
CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
CC syndrome, Alzheimer's disease, tuberculous atherosclerosis, hypercalcaemia, or
CC cerebral palsy. The present DNA sequence encodes a MOL protein of the
CC invention.
XX
XX Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 25; DB 10; Length 1343;
Beat Local Similarity 100.0%; Pred. No. 0.048;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCACTAGTGGCACACCTATGG 25
DB 1090 GAGGCACTAGTGGCACACCTATGG 1066

RESULT 5
ADJ79324/c
ID ADJ79324 atandard; DNA; 1343 BP.
XX
XX ADJ79324;
XX
XX 22-APR-2004 (first entry)
XX
XX NOV11 coding sequence, SEQ ID 21.
XX
XX Cytotoxic; Immunosuppressive; Antiallergic; Antimicrobial; Vasotropic;
KM Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;
KM Ophthalmological; Antiparasitic; Neuroprotective; Nootropic;
KM Anticarcinogenic; Hypotensive; Cardiac; Cerebroprotective;
KM Gene Therapy; NOVX; human; cancer; myelogenous leukemia;
KM congenital neonatal autoimmune thrombocytopenia; immunological disorder;
KM allergy; infection; asthma; lung disease; reproductive disorder;
KM hemangioma; deafness; liver cirrhosis; hepatitis C; gastric disorder;
KM diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;
KM hypertension; stroke; heart failure; chromosome 1; NOV11;
XX chemokine receptor; gene; ds.
XX
XX Homo sapiens.
XX
XX OS

```

XX Key Location/Qualifiers
FH 2..1063
FT CDS /tag=a
FT /product="NOV11"
PT
XX
XX WO2004009635-A2.
XX
XX 29-JAN-2004.
XX
XX 04-OCT-2001; 2001WO-US031292.
XX
XX 20-MAR-2001; 2001US-00813432.
XX
XX (CURA-) CURAGEN CORP.
XX
XX
XX Tampier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS,
PI Vernet CM;
XX
XX WPI; 2004-123380/12.
XX
XX P-PSDB; AD179325.
XX
XX
XX Claim 8; Page 43-44; 158pp; English.
XX
XX The present invention relates to novel NOVX proteins and their coding
CC sequences (AD179304-AD179327). The sequences are useful for the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease associated with the protein, or for diagnosing and treating
CC disorders associated with the NOVX protein, such as cancer, myelogenous
CC leukaemia, congenital neonatal autoimmune thrombocytopenia,
CC immunological disorders, allergy and infection, asthma, lung diseases,
CC reproductive disorders, male and female reproductive diseases,
CC haemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,
CC diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,
CC hypertension, stroke and heart failure. NOV11 represents a new member of
CC the chemokine receptor family and the gene is located on chromosome 1.
XX
XX
XX Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 25; DB 12; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCACTAGTGGCCACACCTATGG 25
Db 1090 GAGGCACTAGTGGCCACACCTATGG 1066

```

```

PI Majumder K;
XX
XX WPI; 2004-268835/25.
XX
XX P-PSDB; ADO56004.
XX
XX
XX Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic
PT nephropathy, acute pancreatitis, strokes and multiple sclerosis.
XX
XX Disclosure; Page 28; 87pp; English.
XX
XX The invention relates to novel isolated NOVX nucleic acids and encoded
CC polypeptides. The nucleic acids, polypeptides and antibodies raised
CC against the polypeptides are useful for preventing or treating diseases
CC associated with aberrant NOVX expression or activity e.g., cancer,
CC obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple
CC sclerosis. The present sequence represents a NOVX nucleic acid of the
CC invention.
XX
XX
XX Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 25; DB 12; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCACTAGTGGCCACACCTATGG 25
Db 1090 GAGGCACTAGTGGCCACACCTATGG 1066

```

RESULT 7
AAD46858/c
ID AAD46858 standard; cDNA; 1526 BP.
AC AAD46858;
XX
XX 27-JAN-2003 (first entry)
XX
XX
XX Human 7TM domain receptor 65507 cDNA.
XX
XX
XX Human, adenosine deaminase; seven transmembrane domain receptor; cancer;
XX 7TM; glycoprotease; immune disorder; IGA deficiency; allergy; arylthmia;
XX rheumatoid arthritis; diabetes; atherosclerosis; cardiovascular disorder;
XX hypertension; ischaemic heart disease; obesity; myocardial infarction;
XX endothelial cell disorder; Graves disease; psoriasis; brain disorder;
XX Parkinson's disease; Alzheimer's disease; haematopoietic disorder;
XX cerebral oedema; metabolic disorder; liver disorder; platelet disorder;
XX chromosome mapping; tissue typing; gene therapy; neuroprotective;
XX cytosolic; anorectic; cardiant; haemostatic; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 139..1200
FT /tag=a
FT /product="Human 65507 protein"
FT /note="This region is specifically claimed as SEQ ID NO:
FT 12 in claim 1 of the specification"
FT
FT sig_peptide 139..306
FT /tag=b
FT mat_peptide 307..1197
FT /tag=c
FT /product="Human mature 65507 protein"
XX
XX
XX WO200274960-A2.
XX
XX 26-SEP-2002.
XX
XX 08-NOV-2001; 2001WO-US051427.
XX
XX
XX 08-NOV-2000; 2000US-0246768P.
XX
XX 08-NOV-2000; 2000US-0246772P.
XX
XX 15-NOV-2000; 2000US-0249185P.
XX
XX (MAJU/) MAJUMDER K.
XX

PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Leiby KR, Kapeller-Libermann R, Glucksmann M;
 XX
 DR WPI, 2002-759898/82.
 DR P-PSDB, AAE29236.
 XX
 PT New 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid molecules,
 PT useful for diagnosing and treating cancer, immune, cardiovascular,
 PT hematopoietic, brain, pain, metabolic, liver or platelet disorders, and
 PT in pharmacogenomics.
 XX
 PS Claim 1, Fig 17, 178pp; English.
 XX
 CC The present invention relates to novel 38650, 28472, 5495, 65507, 81588
 CC or 14354 nucleic acid molecules which encode adenosine deaminase, glyco-
 CC sequence or seven transmembrane domain (7TM) receptor family members.
 CC Sequences of the invention are useful in diagnosing and treating cancer
 CC or aberrant cellular proliferation and/or differentiation (e.g. colon or
 CC lung cancer), immune disorders (e.g. selective IgA deficiency, rheumatoid
 CC arthritis, diabetes, allergy), heart or cardiovascular disorders (e.g.,
 CC hyperension, atherosclerosis, arrhythmias, ischaemic heart disease,
 CC myocardial infarction, thrombus) including endothelial cell disorders
 CC (e.g. psoriasis, Grave's disease), hematopoietic disorders, brain
 CC disorders (e.g. cerebral oedema, Parkinson's or Alzheimer's disease),
 CC pain and metabolic disorders (e.g. obesity), liver disorders or platelet
 CC disorders. They are also useful in screening assays, predictive medicine
 CC (e.g. diagnostic assays, prognostic assays, monitoring clinical trials
 CC and pharmacogenetics) and prophylactic and therapeutic methods. The
 CC nucleic acids may also be used in chromosome mapping, tissue typing and
 CC forensic biology and as surrogate markers. Sequences of the invention are
 CC also used in gene therapy. The present sequence is human 7TM domain
 CC receptor 65507 cDNA
 XX
 SQ Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
 Query Match 100.0%; Score 25; DB 6; Length 1526;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGGCACTAGTTCGCACACCTATGG 25
 Db 1227 GAGGCACTAGTTCGCACACCTATGG 1203

RESULT 8
 ACA60889/c
 ID ACA60889 standard; cDNA; 1526 BP.
 XX
 AC ACA60889;
 XX
 DT 08-JUL-2003 (first entry)
 XX
 DE Human cDNA encoding 65507 protein.
 XX
 KW Human; 88; gene; cancer; aberrant cellular proliferation;
 KW differentiation; immune disorders; heart disorder; brain disorder;
 KW cardiovascular disorder; endothelial cell disorder; pain disorder;
 KW hematopoietic disorder; blood vessel disorder; metabolic disorder;
 KW liver disorder; platelet disorder.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 139..1200
 PT /*tag= a
 FT /product= "Protein 65507"
 FT /note= "This CDS is specifically claimed in claim 1"
 XX
 PN US2003009017-A1.
 XX
 PD 09-JAN-2003.
 XX

PF 08-NOV-2001; 2001US-00012140.
 XX
 XX 08-NOV-2000; 2000US-0246768P.
 PR 08-NOV-2000; 2000US-0246772P.
 PR 15-NOV-2000; 2000US-0249185P.
 XX
 PA (LEIBY/) LEIBY R R.
 PA (KAPF/) KAPPELLER-LIBERMANN R.
 PA (GLUC/) GLUCKSMANN M A.
 XX
 PI Leiby KR, Kapeller-Libermann R, Glucksmann MA;
 XX
 DR WPI, 2003-428888/40.
 DR P-PSDB; ABU09571.
 XX
 PT New isolated 38650, 28472, 5495, 65507, 81588 and 14354 nucleic acid
 PT molecules, useful for diagnosing, treating cancer, pain, or immune,
 PT heart, endothelial cell, hematopoietic, blood vessel, brain, metabolic
 PT and liver disorders.
 XX
 PS Claim 2, Fig 17, 90pp; English.
 XX
 CC The invention relates to an isolated 38650 (encoding adenosine
 CC deaminase), 28472 (encoding a glycoproteinase), 5495 (encoding a 7
 CC transmembrane receptor), 65507, 81588 or 14354 nucleic acid molecule (or
 CC a sequence which is at least 60% identical to the six nucleic acids or
 CC their open reading frames, fragments of at least 15 nucleotides,
 CC naturally occurring variants, or a DNA insert of the plasmid deposited
 CC with the American Type Culture Collection as Accession No. not defined in
 CC the specification, which encodes the amino acid sequence). Also included
 CC are a host cell containing the nucleic acids (used to produce the
 CC proteins), the encoded proteins, an antibody that selectively binds to
 CC the polypeptide, and identifying a compound that binds to/modulates the
 CC activity of the polypeptide. The nucleic acid molecules, polypeptides and
 CC methods are useful for diagnosing, treating cancer, aberrant cellular
 CC proliferation and/or differentiation, immune disorders, heart disorders,
 CC cardiovascular disorders including endothelial cell disorders,
 CC hematopoietic disorders, blood vessel disorders, brain disorders, pain
 CC and metabolic disorders, liver disorders and platelet disorders (many
 CC examples of these disorders are given in the specification). The present
 CC sequence is the Human cDNA 65507
 XX
 SQ Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
 Query Match 100.0%; Score 25; DB 8; Length 1526;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGGCACTAGTTCGCACACCTATGG 25
 Db 1227 GAGGCACTAGTTCGCACACCTATGG 1203

RESULT 9
 ABS57024/c
 ID ABS57024 standard; cDNA; 1526 BP.
 XX
 AC ABS57024;
 XX
 DT 30-JAN-2003 (first entry)
 XX
 DE cDNA encoding novel human 7 transmembrane receptor 65507.
 XX
 KW Cancer; aberrant cell proliferation; aberrant cell differentiation;
 KW breast cancer; ovarian cancer; prostate cancer; colon cancer;
 KW lung cancer; immune disorder; heart disorder; cardiovascular disorder;
 KW endothelial disorder; hematopoietic disorder; blood vessel disorder;
 KW brain disorder; pain; metabolic disorder; liver disorder; diabetes;
 KW platelet disorder; carcinoma; sarcoma; leukaemia; Hodgkin's disease;
 KW autoimmune disorder; hypertension; atherosclerosis; heart failure;
 KW myocardial infarction; ischaemic heart disease; Crohn's disease;
 KW Grave's disease; Kawasaki syndrome; Raynaud's disease; aneurysm;
 KW cerebral ischaemia; peripheral neuropathy; Alzheimer's disease;

KW Parkinson's disease; anorexia nervosa; cachexia; gene; ss;
KM 7 transmembrane receptor.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 139..1200
FT /tag= a
FT /product= "7 transmembrane receptor 65507"
FT /note= "Specifically claimed in claim 1"
XX
PN WO200277233-A2.
XX
PD 03-OCT-2002.
XX
PF 08-NOV-2001; 2001WO-US046724.
XX
PR 08-NOV-2000; 2000US-0246768P.
PR 08-NOV-2000; 2000US-0246772P.
PR 15-NOV-2000; 2000US-0249185P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Leiby KR, Kapeller-Libermann R, Glucksmann M;
XX WPI, 2003-029938/02.
XX P-PSDB; ABB71163.
XX
PT New adenosine deaminase, glycoprotease and seven transmembrane domain
PT nucleic acids and polypeptides, designated 38650, 28472, 5495, 65507,
PT 81588 and 14354, useful for treating e.g. leukemias, Hodgkin's disease or
PT hyperextension.
XX
PS Claim 2; Fig 17; 178pp; English.
XX
CC The invention describes isolated 38650, 28472, 5495, 65507, 81588 and
CC 14354 nucleic acid molecules (I) and their encoded polypeptides (II). The
CC 38650 nucleic acid molecule comprises a sequence encoding adenosine
CC deaminase. The 28472 nucleic acid molecule comprises a sequence encoding
CC a human glycoprotease. The 5495, 65507, 81588 and 14354 comprise
CC sequences that encode a human seven transmembrane domain (7TM). The
CC 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid and polypeptide
CC sequences are useful for diagnosing, preventing or treating a subject
CC with or at risk of developing a disorder, e.g. cancer or aberrant
CC cellular proliferation and/or differentiation (e.g. breast, ovarian,
CC prostate, colon or lung cancer), immune disorders, heart disorders,
CC cardiovascular disorders, endothelial disorders, hematopoietic disorders,
CC blood vessel disorders, brain disorders, pain and metabolic disorders,
CC liver disorders or platelet disorders. These disorders include carcinoma,
CC sarcoma, leukemia, Hodgkin's disease, autoimmune disorders,
CC hypertension, atherosclerosis, heart failure, myocardial infarction,
CC ischemic heart disease, Crohn's disease, Grave's disease, Kawasaki
CC syndrome, Reynaud's disease, aneurysm, cerebral ischaemia, peripheral
CC neuropathy, Alzheimer's disease, Parkinson's disease, anorexia nervosa,
CC cachexia or diabetes. This sequence encodes the novel human 7
CC transmembrane receptor 65507
XX
SQ Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
Query Match 100.0%; Score 25; DB 10; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCGTAGTTCACACCTATGG 25
Db 1227 GAGGCGTAGTTCACACCTATGG 1203

RESULT 10
ABN84272/c
ID ABN84272 standard; cDNA, 1826 BP.
XX
AC ABN84272;

XX 23-SEP-2002 (first entry)
XX
XX Human chemokine-like receptor cDNA.
XX
XX Chemokine-like receptor; G-protein coupled receptor; receptor; human;
XX HIV infection; cardiovascular disease; asthma;
XX chronic obstructive pulmonary disease; cardiac; antidiabetic;
XX vasotropic; hypotensive; antiarrhythmic; thrombolytic; anticoagulant;
XX antiinflammatory; antiallergic; immunomodulator; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PD 20-JUN-2002.
XX
PF 12-DEC-2001; 2001WO-BP014571.
XX
PR 14-DEC-2000; 2000US-0255150P.
PR 02-APR-2001; 2001US-0280110P.
PR 21-JUN-2001; 2001US-0299474P.
XX
PA (FAR) BAYER AG.
XX
PI Smolyar A, Zhu Z, Encinas J, Watanabe S, Okigami H;
XX WPI, 2002-547858/58.
XX P-PSDB; ABB79518.
XX
PT New isolated polynucleotide encoding a chemokine-like receptor
PT polypeptide for treating e.g. asthma, myocardial infarction, human
PT immunodeficiency virus infection, or chronic obstructive pulmonary
PT disease.
XX
XX Disclosure; Fig 6; 114pp; English.
XX
CC The present sequence is that of cDNA encoding a novel human chemokine-
CC like receptor (see AMN79518) of 356 amino acids (41.4 kDa) and having 7
CC putative transmembrane domains, consistent with the structure of a G-
CC protein coupled receptor. Its closest human homologue is C-C chemokine
CC receptor 3. The novel receptor is expressed at low levels in most
CC tissues. It is expressed at a high level in phytohemagglutinin-
CC stimulated CD8+ cells, but in none of the other immune cells tested. It
CC may act as a receptor of chemottractant molecules on activated
CC lymphocytes and be involved in cell trafficking and homing to sites of
CC infection, inflammation or tissue injury. Regulation of activity of the
CC novel receptor can therefore be used to treat cardiovascular,
CC immunological and inflammatory diseases, including asthma and chronic
CC obstructive pulmonary disease (COPD). The receptor may also be a target
CC for viruses that reside in the nervous system. Regulating the binding of
CC ligands, e.g. chemottractant molecules or virus particles, to the
CC receptor can therefore be used to modulate the immune response to inhibit
CC viral infections, including HIV infection. A claimed method of reducing
CC (preferably of the receptor involves contacting a cell with a reagent
CC product (preferably RNA or a polypeptide) encoded by a polynucleotide
CC encoding the human chemokine-like receptor in vivo or in vitro. A claimed
CC method of treating a chemokine-like receptor dysfunction related disease
CC selected from HIV infection, a cardiovascular disorder, asthma or COPD
CC uses a reagent that modulates a function of the receptor
XX
SQ Sequence 1826 BP; 452 A; 510 C; 371 G; 493 T; 0 U; 0 Other;
Query Match 100.0%; Score 25; DB 6; Length 1826;
Best Local Similarity 100.0%; Pred. No. 0.05;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTGCACACTATGG 25
 |||||
 Db 1132 GAGGCACTAGTGCACACTATGG 1108

RESULT 11
 ID ADO28955/c
 ID ADO28955 standard, cDNA; 2273 BP.

XX ADO28955;
 XX
 DT 29-JUL-2004 (first entry)

XX Human novel GPCR PGR3 polynucleotide, SEQ ID NO:54.

XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
 transgenic mouse; neurological disorder; adrenal gland disorder;
 colon disorder; intestinal disorder; cardiovascular disorder;
 muscular disorder; blood disorder; immune disorder; bone disorder;
 joint disorder; metabolic disorder; nutritive disorder; cancer;
 kidney disorder; liver disorder; lung disorder; breast disorder;
 ovary disorder; uterus disorder; prostate disorder; testis disorder;
 skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
 cytostatic; antiinflammatory; vasotropic; antidiagonal; antiserhythmic;
 CNS; central nervous system; respiratory; antidiarrheal; antidiabetic;
 virologic; hepatotropic; antibacterial; antianemic; antiseborrheic;
 dermatologic; antitumor; antithyroid; antiallergic; anorectic;
 immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 gene; ss.

XX Homo sapiens.
 OS
 PN MO200404000-A2.
 XX
 PD 13-MAY-2004.

XX 09-SEP-2003; 2003WO-US028226.
 PF
 XX 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.

XX (PRIM-) PRIMAL INC.
 PA
 XX Galenaris GA, Bergmann JE, Gregorov A, Hohmann J, Li F,
 PI Madisen L, McIlwain KL, Pavlova KM, Vasiliadis D, Zeng H;
 XX WPI; 2004-350329/36.
 DR P-PADB; ADO28954.
 XX

XX Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX

XX Claim 13; SEQ ID NO 54; 542bp; English.

XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of

CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anemia or leukemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancer). The present sequence represents a GPCR-encoding
 CC nucleic acid of the invention. Note: The full sequence data for this
 CC patent did not form part of the printed specification; those sequences
 CC not shown were obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2273 BP; 503 A; 691 C; 549 G; 530 T; 0 U; 0 Other;

XX

XX Query Match 100.0%; Score 25; DB 12; Length 2273;
 Best Local Similarity 100.0%; Pred. No. 0.052;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTGCACACTATGG 25
 |||||
 Db 1541 GAGGCACTAGTGCACACTATGG 1517

RESULT 12
 ID AAH57320/c
 ID AAH57320 standard, cDNA; 295 BP.

XX AAH57320;
 XX
 DT 10-SEP-2001 (first entry)

XX Human pancreas specific cDNA sequence SEQ ID NO:160.

XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
 KW liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
 KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
 KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

XX Homo sapiens.
 OS
 XX
 PN WO200132927-A2.
 XX
 PD 10-MAY-2001.

XX 02-NOV-2000; 2000WO-US030396.
 PF
 XX 04-NOV-1999; 99US-0163508P.

XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Sornasse T, Sellhammer JJ, Watson GA;
 PI WPI; 2001-291057/30.
 DR
 XX

XX New cell and tissue specific polynucleotides useful for diagnosis,
 PT prognosis or monitoring of treatments for disorders where the gene is
 PT associated with a cancer, immunopathology or neuropathology.
 XX

XX Claim 1; Page 125; 327bp; English.

XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
 CC sequences (I). (I) can have cytostatic, immunomodulatory and
 CC neuroprotective activities, and can be used in gene therapy. (I) and
 CC proteins (II) encoded by them are used in high throughput screening
 CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
 CC mimetics, peptides, proteins, agonists, antagonists, antibodies or their
 CC fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical

```
CC agents. Expression of (I) in a sample indicates the differentiation of
CC embryonic stem cells into a tissue selected from brain, heart, kidney,
CC liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
CC to produce an expression profile that defines a metabolic or
CC developmental process, treatment, condition, disease or disorder. The
CC gene profile can be used for diagnosis, prognosis or monitoring of
CC treatments and for investigating a predisposition to a disorder where the
CC gene is associated with a cancer, immunopathology or neuropathology
XX
SQ Sequence 295 BP; 59 A; 82 C; 84 G; 60 T; 0 U; 10 other;
Query Match 72.8%; Score 18.2; DB 4; Length 295;
Best Local Similarity 80.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 GAGGCGATGATGCGACACTGATGG 25
DB 52 GANACGAGTGTGCGACTCCATG 28
RESULT 13
ABK49450
ID ABK49450 standard; DNA; 147309 BP.
XX AC ABK49450;
XX 15-JUN-2002 (first entry)
DE Human transporter genomic DNA sequence.
XX
KW Human; transporter protein; therapeutic target; query sequence;
KW database search; single nucleotide polymorphism; SNP; chromosome 5; gene;
XX ds.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FH exon 2002..2778
FT /*tag= a
FT /number= 1
FT intron 2779..43601
FT /*tag= b
FT /number= 1
FT variation replace(3356,T)
FT /*tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(3396,C)
FT /*tag= w
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(3493,A)
FT /*tag= x
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(3778..3780,GC)
FT /*tag= y
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(4063,G)
FT /*tag= z
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(4628,G)
FT /*tag= aa
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(4702,A)
FT /*tag= ab
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(5235,C)
FT /*tag= ac
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(5638,A)
FT /*tag= ad
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(6248,C)
FT /*tag= ae
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(21708,G)
```

```

FT      /*tag= bd
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(2641,A)
FT      /*tag= bd
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(24758,G)
FT      /*tag= bf
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(25059,C)
FT      /*tag= bg
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26084, .26086, CAA)
FT      /*tag= bh
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26318,G)
FT      /*tag= bi
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(27203, .27205, CTTG)
FT      /*tag= bj
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(27291, .27293, TRG)
FT      /*tag= bk
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(27403, .27405, AYT)
FT      /*tag= bl
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28025,C)
FT      /*tag= bm
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28389,G)
FT      /*tag= bn
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28587,G)
FT      /*tag= bo
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(29581,A)
FT      /*tag= bp
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28727,G)
FT      /*tag= bq
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(30314,A)
FT      /*tag= br
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(30888,G)
FT      /*tag= bs
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(32895,G)
FT      /*tag= bt
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(33150,A)
FT      /*tag= bu
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(33222,T)
FT      /*tag= bv
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(36395, .36397, ATA)
FT      /*tag= bw
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(37006,A)
FT      /*tag= bx
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(37919,A)
FT      /*tag= by
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(38891,T)
FT      /*tag= bz
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(39293,G)
FT      /*tag= ca
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(39933,G)
FT      /*tag= cb

```

```

FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(39968,A)
FT      /*tag= cc

Query Match
Best Local Similarity 72.8%; Score 18.2; DB 6; Length 147309;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GAGCAGTACTTGCCACACTAT 23
Db      122948 GAGCAGTACTTGCCACTCTAT 122970

RESULT 14
AA194220/C
ID      AA194220 standard; cDNA; 879 BP.
XX
AC      AA194220;
XX
DT      13-NOV-2001 (first entry)
XX
DE      Human neuroblastoma expressed polynucleotide SEQ ID NO 295.
XX
KW      Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS      Homo sapiens.
XX
W0200166719-A1.
XX
PD      13-SEP-2001.
XX
PE      02-MAR-2001; 2001MO-JP001629.
XX
PR      07-MAR-2000; 2000JP-00159195.
XX
PA      (CHIB-) CHIBA PREFECTURE.
PA      (HISM ) HISAMITSU PHARM CO LTD.
XX
PI      Nakagawara A;
XX
DR      WPI; 2001-565584/63.
XX
PT      Nucleic acids originating in gene expressed in human neuroblastoma,
PT      useful as probe or primer in diagnosing prognosis of human neuroblastoma,
PT      malignancy and susceptibility indicator or tumor marker for anti-cancer
PT      agents.
XX
PS      Claim 1; Page 258; 2979pp; Japanese.
XX
CC      The invention relates to novel genes (AA193926-AA197963) expressed in
CC      human neuroblastoma. The nucleic acids are applicable as a probe or
CC      primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC      susceptibility indicators or tumor markers for anti-cancer agents. The
CC      gene information for diagnosing prognosis is related to factors similar
CC      to that for N-myc and TrkA genes
XX
SQ      Sequence 879 BP; 200 A; 226 C; 208 G; 216 T; 0 U; 29 Other;

Query Match
Best Local Similarity 70.4%; Score 17.6; DB 4; Length 879;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 AGCAGTAGTTGCCACACTATCG 25
Db      441 AGCAGTAGTTGCCACACTCTTGG 418

RESULT 15
ADO60492
ID      ADO60492 standard; DNA; 1376 BP.
XX
AC      ADO60492;
XX

```

DT 15-JUN-2004 (first entry)
 XX Chlamydomonas ascorbate peroxidase gene SeqID1.
 DE
 XX plant; grass family; cold-resistance; active oxygen elimination;
 KW ascorbate peroxidase; Chlamydomonas; chloroplast transfer sequence;
 KW translocated; chloroplast;
 KW thylakoid membrane proximity transfer sequence; food crop;
 KW cold region cultivation; gene; de.
 XX
 OS Chlamydomonas.
 XX
 FH Key Location/Qualifiers
 FT CDS 32..991
 FT /*tag= a
 FT /product= "Chlamydomonas ascorbate peroxidase"
 XX
 PN JP2004105136-A.
 PD 08-APR-2004.
 XX
 PF 20-SEP-2002; 2002JP-00274942.
 XX
 PR 20-SEP-2002; 2002JP-00274942.
 XX
 PA (SHOK-) SHOKUBUTSU KOGAKU KK.
 PA (KANT) KANSAI DENRYOKU KK.
 XX
 DR WPI; 2004-289542/27.
 DR P-PSDB; ADO60493.
 XX
 PT Novel plants of grass family e.g. rice plants with increased cold-
 PT resistance obtained by reinforcing the expression of gene encoding enzyme
 PT e.g. ascorbate peroxidase, involved in elimination of active oxygen.
 XX
 PS Claim 5; SEQ ID NO 1; 27pp; Japanese.
 XX
 CC This invention relates to novel plants of the grass family whose cold-
 CC resistance property is reinforced by reinforcing the expression of a gene
 CC encoding an enzyme involved in elimination of active oxygen. The plant is
 CC preferably rice and the gene is ascorbate peroxidase gene which is
 CC derived from Chlamydomonas. The ascorbate peroxidase gene comprises a
 CC chloroplast transfer sequence, which allows the gene product to be
 CC translocated to chloroplast. The ascorbate peroxidase gene also comprises
 CC an added thylakoid membrane proximity transfer sequence, which allows the
 CC gene product to translocate in the vicinity of chloroplast. The invention
 CC is useful as a food crop suitable for cultivation in cold regions. The
 CC plants have increased cold-resistance. The present sequence is that of
 CC the Chlamydomonas ascorbate peroxidase gene which is used during the
 CC creation of the plants of the invention.
 CC
 SQ Sequence 1376 BP; 224 A; 473 C; 469 G; 210 T; 0 U; 0 Other;
 Query Match 70.4%; Score 17.6; DB 12; Length 1376;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GAGGCAAGTAGTCCACACCTATG 24
 DB 1273 GCGGCAACAGTAGCCACACCTATG 1296

Search completed: December 14, 2004, 22:43:57
 Job time : 13.7084 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:27.19 : Search time 2.10302 Seconds
(without alignments)
8449.603 Million cell updates/sec

Title: US-09-995-225B-42

Sequence: 1 gagcgacgtatgtccacacactatg 25

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	17.4	69.6	350	4 US-09-513-999C-12278	Sequence 12278, A
C 2	17.2	68.8	381	4 US-09-513-999C-13199	Sequence 13199, A
C 3	16.8	67.2	62804	4 US-09-800-960-3	Sequence 3, Appl1
C 4	16.8	67.2	62804	4 US-10-096-960-3	Sequence 3, Appl1
5	16.6	66.4	795	4 US-09-270-767-1106	Sequence 1106, Ap
6	16.6	66.4	795	4 US-09-270-767-16388	Sequence 16388, A
C 7	16.6	66.4	2379	4 US-09-205-258-175	Sequence 175, App
8	16.2	64.8	46	1 US-08-171-389-311	Sequence 311, App
9	16.2	64.8	46	1 US-08-123-936-311	Sequence 311, App
10	16.2	64.8	46	2 US-08-475-228A-311	Sequence 311, App
11	16.2	64.8	46	3 US-08-482-080A-311	Sequence 311, App
12	16.2	64.8	46	3 US-09-354-947-311	Sequence 311, App
13	16.2	64.8	46	5 PCT-US93-12388-311	Sequence 311, App
14	16.2	64.8	50	4 US-08-956-171E-5107	Sequence 5107, Ap
15	16.2	64.8	50	4 US-08-781-986A-5107	Sequence 5107, Ap
16	16.2	64.8	223	4 US-09-032-297A-12	Sequence 12, Appl
17	16.2	64.8	223	4 US-09-229-151C-6	Sequence 6, Appl1
18	16.2	64.8	335	4 US-08-956-171E-4200	Sequence 4200, Ap
19	16.2	64.8	335	4 US-08-781-986A-4200	Sequence 4200, Ap
20	16.2	64.8	399	4 US-08-956-171E-4338	Sequence 4338, Ap
21	16.2	64.8	399	4 US-08-781-986A-4338	Sequence 4338, Ap
22	16.2	64.8	463	4 US-08-956-171E-3850	Sequence 3850, Ap
23	16.2	64.8	463	4 US-08-781-986A-3850	Sequence 3850, Ap
C 24	16.2	64.8	468	4 US-09-270-767-14303	Sequence 14303, A
C 25	16.2	64.8	657	4 US-08-961-527-314	Sequence 314, App
26	16.2	64.8	904	4 US-09-032-297A-4	Sequence 4, Appl1
27	16.2	64.8	1008	4 US-09-032-297A-7	Sequence 7, Appl1

28	16.2	64.8	1096	4 US-09-032-297A-2	Sequence 2, Appl1
29	16.2	64.8	1096	4 US-09-229-151C-2	Sequence 2, Appl1
C 30	16.2	64.8	1101	4 US-09-370-767-10150	Sequence 10150, A
31	16.2	64.8	1178	4 US-09-032-297A-1	Sequence 1, Appl1
C 32	16.2	64.8	1178	4 US-09-229-151C-1	Sequence 1, Appl1
C 33	16.2	64.8	1179	4 US-09-248-796A-2696	Sequence 2696, Ap
34	16.2	64.8	1357	3 US-09-448-176-9	Sequence 9, Appl1
35	16.2	64.8	1357	3 US-09-448-176-10	Sequence 10, Appl
36	16.2	64.8	1357	3 US-09-448-176-11	Sequence 11, Appl
37	16.2	64.8	1490	4 US-09-032-297A-5	Sequence 5, Appl1
38	16.2	64.8	1587	4 US-09-032-297A-8	Sequence 8, Appl1
39	16.2	64.8	1789	4 US-09-032-297A-6	Sequence 6, Appl1
40	16.2	64.8	1894	4 US-09-032-297A-9	Sequence 9, Appl1
41	16.2	64.8	2270	4 US-09-229-151C-13	Sequence 13, Appl
42	16.2	64.8	2570	4 US-09-229-151C-14	Sequence 14, Appl
C 43	16.2	64.8	3095	3 US-08-434-000A-7	Sequence 7, Appl1
C 44	16.2	64.8	3095	3 US-09-312-157-7	Sequence 7, Appl1
45	16.2	64.8	3634	3 US-09-166-186-1	Sequence 1, Appl1

ALIGNMENTS

```
RESULT 1
US-09-513-999C-12278/c
; Sequence 12278, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122, 487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 12278
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-12278

Query Match      69.6%; Score 17.4; DB 4; Length 350;
Best Local Similarity 94.7%; Pred. No. 22;
Matches 18, Conservative 0, Mismatches 1, Indels 0; Gaps 0;

QY      1 GAGCGAGTAGTTCACAC 19
Db      173 GAGCGAGTAGTTCACAC 155

RESULT 2
US-09-513-999C-13199
; Sequence 13199, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122, 487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 13199
```

```
/ LENGTH: 381
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 318
/ OTHER INFORMATION: k=g or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 319
/ OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-13199

Query Match          68.8%; Score 17.2; DB 4; Length 381;
Best Local Similarity 86.4%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3 GGCAGTAGTGGCCACCTATG 24
        ||||| ||||| ||||| |||||
Db      280 GGCAGAGTTCCACACCTATG 301

RESULT 3
US-09-800-960-3/c
/ Sequence 3, Application US/09800960
/ Patent No. 6387677
/ GENERAL INFORMATION:
/ APPLICANT: YE, Jane et al.
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ FILE REFERENCE: CL001158
/ CURRENT APPLICATION NUMBER: US/09/800,960
/ CURRENT FILING DATE: 2001-03-08
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 62804
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(62804)
/ OTHER INFORMATION: n = A,T,C or G
US-09-800-960-3

Query Match          67.2%; Score 16.8; DB 3; Length 62804;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AGCAGTAGTGGCCACCT 21
        ||||| ||||| ||||| |||||
Db      43074 AGCAGTAGTGGCCACCT 43055

RESULT 4
US-10-096-960-3/c
/ Sequence 3, Application US/10096960
/ Patent No. 6664085
/ GENERAL INFORMATION:
/ APPLICANT: YE, Jane et al.
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ FILE REFERENCE: CL001158DIV
/ CURRENT APPLICATION NUMBER: US/10/096,960
/ CURRENT FILING DATE: 2002-03-14
/ PRIOR APPLICATION NUMBER: 09/800,960
/ PRIOR FILING DATE: 2001-03-08
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 62804
```

```
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(62804)
/ OTHER INFORMATION: n = A,T,C or G
US-10-096-960-3

Query Match          67.2%; Score 16.8; DB 4; Length 62804;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AGCAGTAGTGGCCACCT 21
        ||||| ||||| ||||| |||||
Db      43074 AGCAGTAGTGGCCACCT 43055

RESULT 5
US-09-270-767-1106
/ Sequence 1106, Application US/09270767
/ Patent No. 6703451
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1106
/ LENGTH: 795
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-1106

Query Match          66.4%; Score 16.6; DB 4; Length 795;
Best Local Similarity 82.6%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      3 GGCAGTAGTGGCCACCTATG 25
        ||||| ||||| ||||| |||||
Db      629 GGCAGTAAGCGCCACACGTATG 651

RESULT 6
US-09-270-767-16388
/ Sequence 16388, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 16388
/ LENGTH: 795
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-16388

Query Match          66.4%; Score 16.6; DB 4; Length 795;
Best Local Similarity 82.6%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      3 GGCAGTAGTGGCCACCTATG 25
        ||||| ||||| ||||| |||||
Db      629 GGCAGTAAGCGCCACACGTATG 651

RESULT 7
US-09-205-258-175/c
```

; Sequence 175, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 2379
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1881)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-205-258-175

Query Match 66.4%; Score 16.6; DB 4; Length 2379;
Best Local Similarity 76.0%; Pred. No. 85;
Matches 19; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GAGCGAGTAGTGGCCACACCTATG 25
Db 1699 GAGCGTAAGCTGCCACACCGAKG 1675

RESULT 8
US-08-171-389-311
; Sequence 311, Application US/0817389
; Patent No. 5578444
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/171,389
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/996,783
 FILING DATE: 23-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/723,618
 FILING DATE: 27-JUN-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/081,070
 FILING DATE: 22-JUN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 311:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 46 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
 INDIVIDUAL ISOLATE: (TNF-alpha)
 US-08-171-389-311

```

Query Match Similarity      64.8%; Score 16.2; DB 1; Length 46;
Beet Local Similarity      85.7%; Pred. No. 54;
Matches      18; Conservative      0; Mismatches      3; Indels      0; Gaps      0

QY      2 AGGCAGTAGTGGCACACCTA 22
      ||||| ||||| ||||| |||||
DB      26 AGGCAGTTGTTGGCACACCCA 46

RESULT 9
US-08-123-936-311
; Sequence 311, Application US/08123936
; Patent No. 5726014
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; TITLE OF INVENTION: Screening Assay for the Detection of
; NUMBER OF INVENTION: DNA-Binding Molecules
; NUMBER OF SEQUENCES: 640
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,936
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.

```

1 REGISTRATION NUMBER: 33,875
 2 REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
 3 TELECOMMUNICATION INFORMATION:
 4 TELEPHONE: (415) 324-0880
 5 TELEFAX: (415) 324-0960
 6 INFORMATION FOR SEQ ID NO: 311:
 7 SEQUENCE CHARACTERISTICS:
 8 LENGTH: 46 base pairs
 9 TYPE: nucleic acid
 10 STRANDEDNESS: double
 11 TOPOLOGY: linear
 12 MOLECULE TYPE: DNA (genomic)
 13 HYPOTHEITICAL: NO
 14 ORIGINAL SOURCE:
 15 INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
 16 INDIVIDUAL ISOLATE: (TNF-alpha)
 17 US-08-123-936-311

Query Match	64.8%	Score 16.2	DB 1	Length 46
Best Local Similarity	85.7%	Pred. No. 54		
Matches	18	Conservative 0	Mismatches 3	Indels 0
				Gaps 0
Qy	2	AGGCACTAGTGGCACACTTA	22	
db	26	AGGCACTGTGTGGCAACCCCA	46	

RESULT 10
 US-08-475-228A-311
 Sequence 311. Application US/08475228A
 Patent No. 5869241
 GENERAL INFORMATION:
 APPLICANT: Edwards, Cynthia A.
 APPLICANT: Cantor, Charles R.
 APPLICANT: Andrews, Beth M.
 APPLICANT: Turin, Lisa M.
 APPLICANT: Fry, Kirk E.
 TITLE OF INVENTION: Sequence-Directed DNA Binding
 TITLE OF INVENTION: Molecules, Compositions and Methods
 NUMBER OF SEQUENCES: 664
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genelabs Technologies, Inc.
 STREET: 505 Penobscot Drive
 CITY: Redwood City
 STATE: CA
 COUNTRY: USA
 ZIP: 94063
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/475,228A
 FILING DATE: 06-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/123,936
 FILING DATE: 17-SEP-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/996,783
 FILING DATE: 23-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/723,618
 FILING DATE: 27-JUN-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/081,070
 FILING DATE: 22-JUN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Stratford, Carol A.
 REGISTRATION NUMBER: 34,444
 REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 311:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
INDIVIDUAL ISOLATE: (TNF-alpha)
US-08-475-228A-311

Query Match 64.8%; Score 16.2; DB 2; Length 46;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGGACGAGTGTGGCACACCTA 22
Db 26 AGGACGAGTGTGGCACACCTA 46

RESULT 11
US-08-482-080A-311
Sequence 311, Application US/08482080A
Patent No. 6010849
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,080A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/171,389
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Brady, John P.
REGISTRATION NUMBER: 39,118
REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960

INFORMATION FOR SEQ ID NO: 311:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
INDIVIDUAL ISOLATE: (TNF-alpha)
US-08-482-080A-311

Query Match 64.8%; Score 16.2; DB 3; Length 46;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGGACGAGTGTGGCACACCTA 22
Db 26 AGGACGAGTGTGGCACACCTA 46

RESULT 12
US-09-354-947-311
Sequence 311, Application US/09354947
Patent No. 6384208
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/354,947
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/482,080
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/171,389
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Brady, John P.
REGISTRATION NUMBER: 39,118
REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880

TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 311:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
INDIVIDUAL ISOLATE: (TNF-alpha)
US-09-354-947-311

Query Match 64.8%; Score 16.2; DB 3; Length 46;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AGGCAAGTGTGCGCACCTA 22
Db 26 AGGCAAGTGTGCGCACCTA 46

RESULT 13
PCT-US93-12388-311
Sequence 311, Application PC/TUS9312388
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESSER: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12388
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175.41/519PCT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 311:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
INDIVIDUAL ISOLATE: (TNF-alpha)
PCT-US93-12388-311

Query Match 64.8%; Score 16.2; DB 5; Length 46;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AGGCAAGTGTGCGCACCTA 22
Db 26 AGGCAAGTGTGCGCACCTA 46

RESULT 14
US-08-956-171E-5107
Sequence 5107, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunesh
Gill H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA

ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5107:

SEQUENCE CHARACTERISTICS:

LENGTH: 50 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 5107:
US-08-956-171E-5107

Query Match 64.8%; Score 16.2; DB 4; Length 50;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 GCAGTAGTGTGCGCACCTATG 24
Db 7 GCAGTAGTGTGCGCACCTATG 27

RESULT 15
US-08-781-986A-5107
Sequence 5107, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:

APPLICANT: Charles Kunsch
TITLE OF INVENTION: *Saaphylococcus aureus* Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5235
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5107:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-5107

	Query Match	Similarity	Score	DB	length
Best Local	18	85.7%	Pred. No. 55		
Matches	18	Conservative	0	Mismatches	3
				Indels	0
				Gaps	0
Qy	4	GCAGTATGTCACACTTATG	24		
Db	7	GCAGTATGTCACACTTGTG	27		

Search completed: December 15, 2004, 02:56:26
Job time : 3.10302 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using aw model

Run on: December 14, 2004, 22:25:49 : Search time 12.7717 Seconds
(without alignments)
10792.424 Million cell updates/sec

Title: US-09-995-225B-42

Perfect score: 25

Sequence: 1 gaggcagtagtcgcacactatcg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	9	US-09-995-225-42
2	25	100.0	25	10	US-09-995-225-42
3	25	100.0	1202	14	US-10-094-417-1
4	25	100.0	1343	15	US-09-813-432-21
5	25	100.0	1343	15	US-10-174-364-21
6	25	100.0	1343	16	US-10-246-583-21
7	25	100.0	1343	17	US-10-689-832-21
8	25	100.0	1526	14	US-10-012-140-10
9	25	100.0	1826	16	US-10-450-590-6
10	25	100.0	2117	17	US-10-779-104-1
11	19.2	76.8	688	16	US-10-424-599-59902
12	18.2	72.8	1886	17	US-10-437-963-65084

13	18.2	72.8	147309	9	US-09-742-312-3	Sequence 3, Appl1
14	18.2	72.8	147309	15	US-10-436-185-3	Sequence 3, Appl1
15	17.8	71.2	275	18	US-10-425-115-19752	Sequence 19752, A
16	17.8	71.2	787	17	US-10-767-701-14254	Sequence 14254, A
17	17.6	70.4	519	13	US-10-027-632-94559	Sequence 94559, A
18	17.6	70.4	519	13	US-10-027-632-94559	Sequence 94560, A
19	17.6	70.4	519	13	US-10-027-632-305465	Sequence 305465, A
20	17.6	70.4	519	13	US-10-027-632-305465	Sequence 305466, A
21	17.6	70.4	519	15	US-10-027-632-94559	Sequence 94559, A
22	17.6	70.4	519	15	US-10-027-632-94560	Sequence 94560, A
23	17.6	70.4	519	15	US-10-027-632-94565	Sequence 305465, A
24	17.6	70.4	519	15	US-10-027-632-305466	Sequence 305466, A
25	17.6	70.4	542	16	US-10-424-599-128896	Sequence 128896, A
26	17.6	70.4	152501	17	US-10-116-231-4	Sequence 4, Appl1
27	17.4	69.6	208	10	US-09-814-353-339	Sequence 339, App
28	17.4	69.6	208	10	US-09-814-353-6722	Sequence 6722, App
29	17.4	69.6	248	10	US-09-814-353-19298	Sequence 19298, A
30	17.4	69.6	527	10	US-09-814-353-13107	Sequence 13107, A
31	17.4	69.6	674	9	US-09-764-570-272	Sequence 272, App
32	17.4	69.6	674	14	US-10-125-540-272	Sequence 272, App
33	17.4	69.6	1596	15	US-10-371-264-2	Sequence 2, Appl1
34	17.4	69.6	1596	15	US-10-371-099-2	Sequence 2, Appl1
35	17.4	69.6	1596	16	US-10-371-122-2	Sequence 2, Appl1
36	17.4	69.6	1596	16	US-10-373-567-2	Sequence 2, Appl1
37	17.4	69.6	1596	16	US-10-628-088-2	Sequence 2, Appl1
38	17.4	69.6	1666	15	US-10-371-264-3	Sequence 3, Appl1
39	17.4	69.6	1666	15	US-10-371-099-3	Sequence 3, Appl1
40	17.4	69.6	1666	16	US-10-371-122-3	Sequence 3, Appl1
41	17.4	69.6	1666	16	US-10-373-567-3	Sequence 3, Appl1
42	17.4	69.6	1666	16	US-10-628-088-3	Sequence 3, Appl1
43	17.4	69.6	1860	15	US-10-371-264-5	Sequence 5, Appl1
44	17.4	69.6	1860	15	US-10-371-099-5	Sequence 5, Appl1
45	17.4	69.6	1860	16	US-10-371-122-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1

US-09-995-225-42

Sequence 42, Application US/09995225

Publication No. US20020193584A1

GENERAL INFORMATION:

APPLICANT: Chen, Ruoping

APPLICANT: Chu, Zhi Liang

APPLICANT: Dang, Huong T.

APPLICANT: Lowitz, Kevin P.

APPLICANT: Pride, Cameron

TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human

FILE REFERENCE: AREN-0308

CURRENT APPLICATION NUMBER: US/09/995, 225

PRIOR FILING DATE: 2001-11-26

PRIOR APPLICATION NUMBER: 09/170, 496

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: PCT/US99/23938

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: 60/253, 404

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/255, 366

PRIOR FILING DATE: 2000-12-12

PRIOR APPLICATION NUMBER: 60/270, 286

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282, 365

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/270, 266

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282, 032

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282, 358

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282, 356

PRIOR FILING DATE: 2001-04-06

```
/ PRIOR APPLICATION NUMBER: 60/290,917
/ PRIOR FILING DATE: 2001-05-14
/ PRIOR APPLICATION NUMBER: 60/309,208
/ PRIOR FILING DATE: 2001-07-31
/ NUMBER OF SEQ ID NOS: 67
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 42
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: No. US20020193584A1e1 Sequence
US-09-995-225-42
```

```
Query Match          100.0%; Score 25; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GAGGCAGTAGTGGCCACACCTATGG 25
        |||||||
Db       1 GAGGCAGTAGTGGCCACACCTATGG 25
```

```
RESULT 2
US-09-995-225-42
/ Sequence 42, Application US/09995225
/ Publication No. US20030139588A9
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Ruoping
/ APPLICANT: Chu, Zhi Liang
/ APPLICANT: Dang, Huong T.
/ APPLICANT: Lowitz, Kevin P.
/ APPLICANT: Pride, Cameron
/ TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
/ TITLE OF INVENTION: Receptors
/ FILE REFERENCE: ALEN-0308
/ CURRENT APPLICATION NUMBER: US/09/995,225
/ PRIOR FILING DATE: 2001-11-26
/ PRIOR APPLICATION NUMBER: 09/170,496
/ PRIOR FILING DATE: 1998-10-13
/ PRIOR APPLICATION NUMBER: PCT/US99/23938
/ PRIOR FILING DATE: 1998-10-13
/ PRIOR APPLICATION NUMBER: 60/253,404
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/255,366
/ PRIOR FILING DATE: 2000-12-12
/ PRIOR APPLICATION NUMBER: 60/270,286
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/282,365
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/270,266
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/282,032
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,358
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,356
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/290,917
/ PRIOR FILING DATE: 2001-05-14
/ PRIOR APPLICATION NUMBER: 60/309,208
/ PRIOR FILING DATE: 2001-07-31
/ NUMBER OF SEQ ID NOS: 67
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 42
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: No. US20030139588A9e1 Sequence
US-09-995-225-42
```

Query Match 100.0%; Score 25; DB 10; Length 25;

Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 GAGGCAGTAGTGGCCACACCTATGG 25
        |||||||
Db       1 GAGGCAGTAGTGGCCACACCTATGG 25
```

```
RESULT 3
US-10-094-417-1/C
/ Sequence 1, Application US/10094417
/ Publication No. US20030045685A1
/ GENERAL INFORMATION:
/ APPLICANT: Tian, Hui
/ APPLICANT: Zhao, Jiagang
/ APPLICANT: Chen, Jin-Long
/ APPLICANT: Cutler, Gene
/ APPLICANT: Tularik Inc.
/ TITLE OF INVENTION: No. US20030045685A1e1 Receptors
/ FILE REFERENCE: 018781-008110US
/ CURRENT APPLICATION NUMBER: US/10/094,417
/ CURRENT FILING DATE: 2002-06-10
/ PRIOR APPLICATION NUMBER: US 09/802,803
/ PRIOR FILING DATE: 2001-03-09
/ PRIOR APPLICATION NUMBER: US 60/276,649
/ PRIOR FILING DATE: 2001-03-16
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 1202
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
/ NAME/KEY: CDS
/ LOCATION: (68)..(1129)
/ OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
US-10-094-417-1
```

Query Match 100.0%; Score 25; DB 14; Length 1202;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 GAGGCAGTAGTGGCCACACCTATGG 25
        |||||||
Db       1156 GAGGCAGTAGTGGCCACACCTATGG 1132
```

```
RESULT 4
US-09-813-432-21/C
/ Sequence 21, Application US/09813432
/ Publication No. US20030148485A1
/ GENERAL INFORMATION:
/ APPLICANT: Taupier Jr., Raymond J
/ APPLICANT: Majmuder, Kamud
/ APPLICANT: Spaderma, Steven K
/ APPLICANT: Smithson, Glenda
/ APPLICANT: Mezes, Peter S
/ APPLICANT: Vermet, Corine A. M.
/ TITLE OF INVENTION: No. US20030148485A1e1 Polypeptides and Amino Acids Encoding Same
/ FILE REFERENCE: 15966-729
/ CURRENT APPLICATION NUMBER: US/09/813,432
/ CURRENT FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
```

PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,836
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/193,843
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 1343
TYPE: DNA
ORGANISM: Homo sapiens
US-09-813-432-21

Query Match 100.0%; Score 25; DB 10; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGGACGAGTGTGCGACACCTATGG 25
DB 1090 GAGGACGAGTGTGCGACACCTATGG 1066

RESULT 5

US-10-174-364-21/c
Sequence 21, Application US/10174364
Publication No. US20030216308A1
GENERAL INFORMATION:
APPLICANT: Anderson et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-729CIP2
CURRENT APPLICATION NUMBER: US/10/174,364
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,836
PRIOR FILING DATE: 2000-03-29
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 1343
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(1060)
US-10-174-364-21

Query Match 100.0%; Score 25; DB 15; Length 1343;

Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGGACGAGTGTGCGACACCTATGG 25
DB 1090 GAGGACGAGTGTGCGACACCTATGG 1066

RESULT 6

US-10-246-583-21/c
Sequence 21, Application US/10246583
Publication No. US20040058862A1
GENERAL INFORMATION:
APPLICANT: Majumder
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-729CIP2CON1
CURRENT APPLICATION NUMBER: US/10/246,583
CURRENT FILING DATE: 2002-12-06
PRIOR APPLICATION NUMBER: 10/174,364
PRIOR FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 1343
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(1060)
US-10-246-583-21

Query Match 100.0%; Score 25; DB 16; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGGACGAGTGTGCGACACCTATGG 25
DB 1090 GAGGACGAGTGTGCGACACCTATGG 1066

RESULT 7

US-10-689-832-21/c
Sequence 21, Application US/10689832
Publication No. US20040121380A1
GENERAL INFORMATION:
APPLICANT: Majumder, Kamud
TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
FILE REFERENCE: 15966-729DIV1
CURRENT APPLICATION NUMBER: US/10/689,832
CURRENT FILING DATE: 2003-10-20
PRIOR APPLICATION NUMBER: 09/813,432
PRIOR FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20

PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 1343
TYPE: DNA
ORGANISM: Homo sapiens
US-10-689-832-21

Query Match 100.0%; Score 25; DB 17; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCACTAGTGGCACCTATGG 25
DB 1090 GAGGCACTAGTGGCACCTATGG 1066

RESULT 8
US-10-012-140-10/c
Sequence 10, Application US/10012140
Publication No. US2003009017A1
GENERAL INFORMATION:
APPLICANT: leihey, Kevin R.
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Gluckemann, Maria A.
TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
FILE REFERENCE: THEREOF
CURRENT APPLICATION NUMBER: US/10/012,140
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: 60/246,768
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,772
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,185
PRIOR FILING DATE: 2000-11-15
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 1526
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (139)...(1200)
US-10-012-140-10

Query Match 100.0%; Score 25; DB 14; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCACTAGTGGCACCTATGG 25
DB 1227 GAGGCACTAGTGGCACCTATGG 1203

RESULT 9
US-10-450-590-6/c
Sequence 6, Application US/10450590
Publication No. US20040076985A1
GENERAL INFORMATION:
APPLICANT: Bayer AG
TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
FILE REFERENCE: L10316 Foreign Countries
CURRENT APPLICATION NUMBER: US/10/450,590
CURRENT FILING DATE: 2003-06-16
PRIOR APPLICATION NUMBER: US 60/254,923
PRIOR FILING DATE: 2000-12-14
PRIOR APPLICATION NUMBER: US 60/280,110
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: US 60/299,474
PRIOR FILING DATE: 2001-06-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 1826
TYPE: DNA
ORGANISM: Homo sapiens
US-10-450-590-6

Query Match 100.0%; Score 25; DB 16; Length 1826;
Best Local Similarity 100.0%; Pred. No. 0.063;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCACTAGTGGCACCTATGG 25
DB 1132 GAGGCACTAGTGGCACCTATGG 1108

RESULT 10
US-10-779-104-1/c
Sequence 1, Application US/10779104
Publication No. US20040161799A1
GENERAL INFORMATION:
APPLICANT: Andrew J. Murphy
APPLICANT: Susan Croll-Kalish
TITLE OF INVENTION: KOR3L-like-Proteins and Methods of Modulating KOR3L-Mediated Activ
FILE REFERENCE: REG 1000A
CURRENT APPLICATION NUMBER: US/10/779,104
CURRENT FILING DATE: 2004-02-13
PRIOR APPLICATION NUMBER: 60/447,447
PRIOR FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: 60/495,577
PRIOR FILING DATE: 2003-08-14
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2117
TYPE: DNA
ORGANISM: homo sapiens
US-10-779-104-1

Query Match 100.0%; Score 25; DB 17; Length 2117;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCACTAGTGGCACCTATGG 25
DB 1368 GAGGCACTAGTGGCACCTATGG 1344

RESULT 11
US-10-424-599-59902/c
Sequence 59902, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K

```
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 58902
/ LENGTH: 688
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_25101C.1
US-10-424-599-59902

Query Match          76.8%; Score 19.2; DB 16; Length 688;
Best Local Similarity 87.5%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAGCGACTAGTGGCCACACCTATG 24
DB 353 GAGCGACTAGTGGCCGACCAATG 330

RESULT 12
US-10-437-963-65084/C
/ Sequence 65084, Application US/10437963
/ Publication No. US2004012343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbasuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 65084
/ LENGTH: 1886
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_66167C.1
US-10-437-963-65084

Query Match          72.8%; Score 18.2; DB 17; Length 1886;
Best Local Similarity 87.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAGCGACTAGTGGCCACACCTAT 23
DB 78 GAGCGACTAGTGGCCATAGCTAT 56

RESULT 13
US-09-742-312-3
/ Sequence 3, Application US/09742312
/ Patent No. US20020045166A1
/ GENERAL INFORMATION:
/ APPLICANT: CHANDRAMOULISARAN, Ishwar et al
/ TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
/ FILE REFERENCE: CL000838
/ CURRENT APPLICATION NUMBER: US/09/742,312
/ CURRENT FILING DATE: 2000-12-22
```

```
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 147309
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)...(147309)
/ OTHER INFORMATION: n = A,T,C or G
US-09-742-312-3

Query Match          72.8%; Score 18.2; DB 9; Length 147309;
Best Local Similarity 87.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAGCGACTAGTGGCCACACCTAT 23
DB 122948 GAGCGACTAGTGGCCATTCCTAT 122970

RESULT 14
US-10-436-185-3
/ Sequence 3, Application US/10436185
/ Publication No. US20030180887A1
/ GENERAL INFORMATION:
/ APPLICANT: CHANDRAMOULISARAN, Ishwar et al
/ TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
/ FILE REFERENCE: CL000838CON
/ CURRENT APPLICATION NUMBER: US/10/436,185
/ CURRENT FILING DATE: 2003-05-13
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 147309
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)...(147309)
/ OTHER INFORMATION: n = A,T,C or G
US-10-436-185-3

Query Match          72.8%; Score 18.2; DB 15; Length 147309;
Best Local Similarity 87.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAGCGACTAGTGGCCACACCTAT 23
DB 122948 GAGCGACTAGTGGCCATTCCTAT 122970

RESULT 15
US-10-425-115-19752
/ Sequence 19752, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 19752
/ LENGTH: 275
/ TYPE: DNA
/ ORGANISM: Zea mays
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:17:59, Search time 78.7335 Seconds
(without alignments)
11570.599 Million cell updates/sec

Title: US-09-995-225B-42

Perfect score: 25
Sequence: 1 gggcagtagtcgcacacactatg 25

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1.*
2: gb_est2.*
3: gb_hnc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gsa1.*
9: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	25	100.0	448	AQ225693	Hs_2009 B
2	20.2	80.8	746	CD636851	56003360H
3	19	76.0	590	CL686359	PR10144a
4	19	76.0	808	CL692939	PR1015b F
5	18.8	75.2	172	AW799209	RCO-UM005
6	18.8	75.2	334	AT029980	UT-R-CO-1
7	18.8	75.2	565	CD636852	56003360J
8	18.8	75.2	697	B0616836	UI-H-FH1
9	18.8	75.2	729	CO138524	EST833195
10	18.6	74.4	453	AZ009690	RPCI-23-3
11	18.6	74.4	483	B2660234	B2660234
12	18.6	74.4	489	BT680378	458624 MA
13	18.6	74.4	495	BT260587	BT260587
14	18.6	74.4	504	BK499223	DKF2p779E
15	18.6	74.4	506	AW653346	101999 MA
16	18.6	74.4	543	AV591426	AV591426
17	18.6	74.4	546	BF076557	226150 MA
18	18.6	74.4	557	BG691654	340911 BA
19	18.6	74.4	572	BM030860	495388 MA
20	18.6	74.4	590	BE590190	197189 BA
21	18.6	74.4	590	BI682340	463396 MA
22	18.6	74.4	613	AL781067	AL781067
23	18.6	74.4	619	AL802528	AL802528
24	18.6	74.4	621	BT260398	BT260398

25	18.6	74.4	639	1	AL656915
26	18.6	74.4	640	1	AL876764
27	18.6	74.4	642	1	AL877073
28	18.6	74.4	652	7	CK835565
29	18.6	74.4	661	1	AL629682
30	18.6	74.4	708	7	CK951479
31	18.6	74.4	745	7	CK832878
32	18.6	74.4	774	5	BU238619
33	18.6	74.4	785	5	EX739364
34	18.6	74.4	803	9	CC905917
35	18.4	73.6	294	6	CD850431
36	18.4	73.6	311	6	CD850445
37	18.4	73.6	407	6	CB807057
38	18.4	73.6	579	6	CD849483
39	18.4	73.6	591	6	CD849077
40	18.4	73.6	626	6	CD849065
41	18.4	73.6	670	6	CD848640
42	18.4	73.6	675	6	CD848811
43	18.4	73.6	682	6	CD850036
44	18.4	73.6	683	6	CD850069
45	18.4	73.6	689	6	CD849410

ALIGNMENTS

RESULT 1
AQ225693
LOCUS
DEFINITION
HS 2009 B2 B08 T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2009 Col=16 Row=D, genomic survey
sequence.
AQ225693 GI:3650922
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 448)
Mahatras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahatras GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2009 row: D column: 16
Class: BAC ends
High quality sequence stop: 448.
Location/Qualifiers
1. 448
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2009 Col=16 Row=D"
/sex="male"
/clone_1lb="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-coli DH10B"

ORIGIN
Query Match 100.0%; Score 25; DB 8; Length 448;
Best Local Similarity 100.0%; Pred. No. 0.66;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCACTAGTTCACACCTATG 25
 |||||
 Db 151 GAGGCACTAGTTCACACCTATG 175

RESULT 2
 LOCUS CD636851/c 746 bp mRNA linear EST 12-JAN-2004
 DEFINITION 56003360H1 FLP Homo sapiens cDNA, mRNA sequence.
 ACCESSION CD636851
 VERSION CD636851.1 GI:40285118
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 746)
 AUTHORS Fu, G.K., Wang, Y.T., Yang, J., Au-Young, J. and Stuve, L.L.
 TITLE Circular rapid amplification of cDNA ends for high-throughput
 extension cloning of partial genes
 JOURNAL Genomics 84 (1), 205-210 (2004)
 COMMENT Contact: Fu GK
 Incyte Genomics, Inc.
 3160 Porter Dr., Palo Alto, CA 94304, USA
 Tel: 6508454102
 Email: gfu@incyte.com.

FEATURES
 Source location/Qualifiers
 1..746
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="FLP"
 /note="Vector: pDrive Cloning Vector"

ORIGIN
 Query Match 80.8%; Score 20.2; DB 6; Length 746;
 Best Local Similarity 88.0%; Pred. No. 1.3e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAGGCACTAGTTCACACCTATG 25
 |||||
 Db 658 GTGGCAGAGTTCCACACCTATG 634

RESULT 3
 LOCUS CL686359 590 bp DNA linear GSS 09-JUL-2004
 DEFINITION PR10144a.B01.2 - PR10144a.BR (590) Mixed stage fosmid library of P.
 pacificus var. California Pristionchus pacificus genomic, genomic
 survey sequence.
 ACCESSION CL686359
 VERSION CL686359.1 GI:50194822
 KEYWORDS GSS.
 SOURCE Pristionchus pacificus
 ORGANISM Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 REFERENCE 1 (bases 1 to 590)
 AUTHORS Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.
 TITLE AppADB: an Acedb database for the nematode satellite organism
 JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)
 COMMENT Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.

Seq primer: T7
 Class: fosmid ends.
 Location/Qualifiers
 1..590
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pBifos-5 Fosmid vector"

ORIGIN
 Query Match 76.0%; Score 19; DB 9; Length 590;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCACTAGTTCACAC 19
 |||||
 Db 528 GAGGCACTAGTTCACAC 546

RESULT 4
 LOCUS CL692939 808 bp DNA linear GSS 10-JUL-2004
 DEFINITION PR1015b.F10.2 - PR1015b.BR (808) Mixed stage fosmid library of P.
 pacificus var. California Pristionchus pacificus genomic, genomic
 survey sequence.
 ACCESSION CL692939
 VERSION CL692939.1 GI:50214847
 KEYWORDS GSS.
 SOURCE Pristionchus pacificus
 ORGANISM Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 REFERENCE 1 (bases 1 to 808)
 AUTHORS Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.
 TITLE AppADB: an Acedb database for the nematode satellite organism
 JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)
 COMMENT Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.
 Seq primer: T7
 Class: fosmid ends.
 Location/Qualifiers
 1..808
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pBifos-5 Fosmid vector"

ORIGIN
 Query Match 76.0%; Score 19; DB 9; Length 808;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCACTAGTTCACAC 19
 |||||
 Db 528 GAGGCACTAGTTCACAC 546

RESULT 5
 AW799209

LOCUS AW799209 172 bp mRNA linear EST 16-MAY-2000
 DEFINITION RCO-UM0051-010300-021-c10 UM0051 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW799209
 VERSION AW799209.1 GI:7851079
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 172)
 AUTHORS Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.P., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2a-RCO-UM0051-010300-021-c10&t3=2000-03-01&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 16
 High quality sequence stop: 172.
 FEATURES
 source location/Qualifiers
 1..172
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="UM0051"
 /note="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 ORIGIN
 Query Match 75.2%; Score 18.8; DB 2; Length 172;
 Best Local Similarity 90.9%; Pred. No. 5e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Cyt 4 GCAGTAGTGGCACACCTATGG 25
 |||||||
 Db 134 GCAGTAGTGGCACACCTATGG 155
 |||||||
 RESULT 6 334 bp mRNA linear EST 04-JUN-1999
 LOCUS AI029980
 DEFINITION UI-R-C0-jr-d-11-0-UI-81 UI-R-C0 Rattus norvegicus cDNA clone
 ACCESSION AI029980
 VERSION AI029980.1 GI:3247806
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 334)
 AUTHORS Bonaldo,M.F., Lemmon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Oligo-dt track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.
 Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
 through Research Genetics This clone is also available through the
 I.M.A.G.E. Consortium at LBNL (info@image.lbnl.gov). IMAGE
 ID=1783323
 Seq primer: M13 Forward
 POLYA=No.
 FEATURES
 source location/Qualifiers
 1..334
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C0-jr-d-11-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-C0"
 /note="Vector: pRT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C0 library is a subtracted library derived from the UI-R-A1 and UI-R-B1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UI-R-B1 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dt track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C0) was constructed as follows: PCR amplified cDNA inserts from a pool of UI-R-A1 and UI-R-B1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the pooled UI-R-A1 and UI-R-B1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C0 library. This procedure has been previously described (Bonaldo, Lemmon and Soares, Genome Research 6: 791-806, 1996)"
 ORIGIN
 Query Match 75.2%; Score 18.8; DB 1; Length 334;
 Best Local Similarity 90.9%; Pred. No. 5.5e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Cyt 4 GCAGTAGTGGCACACCTATGG 25
 |||||||
 Db 162 GCAGTAGTGGCACACCTATGG 141
 |||||||
 RESULT 7 565 bp mRNA linear EST 12-JAN-2004
 LOCUS CD636852
 DEFINITION 56003360J1 FLP Homo sapiens cDNA, mRNA sequence.

Query Match	75.2%	Score 18.8;	DB 5;	Length 697;
Best Local Similarity	90.9%	Pred. No. 6e+02;		
Matches	20;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;
Db	1	GAGGCACTAGTTCGCACACCTTA	22	
	288	GAGGCACTGCTTCCACACCTTA	309	
RESULT 9				
LOCUS	COL138524	729 bp	mRNA	linear
DEFINITION	EST833195 Aspergillus flavus Normalized cDNA Expression Library	EST 17-JUN-2004		
ACCESSION	Aspergillus flavus cDNA clone MAFU61 5' end similar to (Q7RG55)			
VERSION	ENSANG00000023647 (Fragment), mRNA sequence.			
KEYWORDS	COL138524			
SOURCE	COL138524.1	GI:48887502		
ORGANISM	Aspergillus flavus			
REFERENCE	Aspergillus flavus			
AUTHORS	Eukaryota Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
DATE	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
COMMENT	1 (bases 1 to 729)			
	Yu, J., Whitelaw, C.A., Nieman, W.C., Bhatnagar, D. and Cleveland, T.E.			
	Aspergillus flavus expressed sequence tags for identification of			
	genes with putative roles in aflatoxin contamination of crops			
	FEMS Microbiol. Lett. (2004) In press			
	Contact: Yu J			
	Food and Feed Safety Research Unit			
	USDA/ARS, Southern Regional Research Center			
	1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA			
	Tel: 504 286 4405			
	Fax: 504 286 4419			
	Email: jinyu@errc.ars.usda.gov			
	Contact Dr. Yu at USDA/ARS SRRC (jinyu@errc.ars.usda.gov) for clone			
	information			
	PCR Primers			
	FORWARD: M13P			
	BACKWARD: M13R			
	Seq primer: M13 Forward			
	POLYA=NO.			
FEATURES				
Source	Location/Qualifiers			
	1..729			
	/organism="Aspergillus flavus"			
	/mol_type="mRNA"			
	/strain="NRRL 3357"			
	/db_xref="taxon:5059"			
	/clone="NAFCU61"			
	/sex="asexual mycelia"			
	/cell_type="mycelia"			
	/dev_stage="developmental stages from 18 to 96 hours"			
	/lab_host="E. coli DH10B T1 resistant cells"			

/clone_1lb="Aspergillus flavus Normalized cDNA Expression Library"
 /note="Vector: pBluescript (SK-) (Stratagene), antibiotic selection marker: Carbenicillin, Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

ORIGIN

Query Match 75.2%; Score 18.8; DB 7; Length 729;
 Best Local Similarity 90.9%; Pred. No. 6e+02; Mismatches 2; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 AGGCAAGTGTGCGACACTAT 23
 Db 652 AGGCAATAGTTGCACACCCAT 673

RESULT 10 453 bp DNA linear GSS 25-FEB-2000
 AZ009690/c RPCI-23-352C3.TV RPCI-23 Mus musculus genomic clone RPCI-23-352C3,
 LOCUS genomic survey sequence.
 DEFINITION
 ACCESSION AZ009690 GI:7085074
 VERSION
 KEYWORDS
 SOURCE GSS.
 ORGANISM Mus musculus (house mouse)

REFERENCE
 AUTHORS Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 453)
 Zhao, S., Niernman, W., Feldblyum, T., Malek, V., Shatsman, S., Akimov, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-23
 TITLE
 JOURNAL
 COMMENT Other_GSS88: RPCI-23-352C3.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_end/mouse/bac_end_intro.html
 Plate: 352 row: C column: 3
 Seg primer: T7
 Class: BAC ends

FEATURES

Source Location/Qualifiers
 1..453
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="CS7BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-352C3"
 /sex="Female"
 /lab_host="DH10B"
 /clone_1lb="RPCI-23"
 /note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Female CS7BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 74.4%; Score 18.6; DB 8; Length 453;
 Best Local Similarity 84.0%; Pred. No. 7e+02; Mismatches 4; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GAGGCAAGTGTGCGACACTATG 25
 Db 365 GAGGCAATGCTTGCACACCATG 341

RESULT 11 483 bp mRNA linear EST 08-APR-2002
 BJ260234/c BJ260234
 LOCUS BJ260234 Y. Ogihara unpublished cDNA library, Wh_h Triticum
 DEFINITION aestivum cDNA clone whn24a16 5', mRNA sequence.
 ACCESSION BJ260234
 VERSION BJ260234.1 GI:20082291
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.
 1 (bases 1 to 483)
 Ogihara, Y. and Murai, K.
 Expressed genes in Triticum aestivum
 Unpublished (2002)
 Contact: Tadasu Shin-1
 Center for Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin1@genie.nig.ac.jp.

FEATURES

Source Location/Qualifiers
 1..483
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="whn24a16"
 /tissue_type="spike at heading date"
 /dev_stage="Feekes' scale 10.5"
 /clone_1lb="Y. Ogihara unpublished cDNA library, Wh_h"

ORIGIN

Query Match 74.4%; Score 18.6; DB 4; Length 483;
 Best Local Similarity 84.0%; Pred. No. 7.1e+02; Mismatches 4; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GAGGCAAGTGTGCGACACTATG 25
 Db 270 GAGGCAATGCTTGCACACCATG 246

RESULT 12 489 bp mRNA linear EST 17-SEP-2001
 BI680378/c BI680378
 LOCUS BI680378 458624 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION
 ACCESSION BI680378
 VERSION BI680378.1 GI:15633289
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 489)
 Smith, T.P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrnenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.M., Rohrer, G.A., Chitko-Mckown, C.G., Pettea, G., Holt, I., Karamycheva, S., Liang, F.,

TITLE
Quackenbush, J. and Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
COMMENT
11282978
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCAGTCACGACG
Plate: 126 row: C column: 19
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
1..489
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_1lb="MARC 1BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI, Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

FEATURES

source

ORIGIN

Query Match 74.4%; Score 18.6; DB 4; Length 489;
Best Local Similarity 84.0%; Pred. No. 7.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GAGGCAAGTAGTGCACACCTATGG 25
|||||
Db 144 GAGGCAAGTAGTGCACCTCCAGTGG 120
|||||

RESULT 13
Bj260587/c 495 bp mRNA linear EST 08-APR-2002
LOCUS Bj260587 Y. Ogihara unpublished cDNA library, wh_h Triticum
DEFINITION aestivum cDNA clone wh25m17 5', mRNA sequence.
ACCESSION Bj260587 GI:20082510
VERSION Bj260587.1
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticaceae; Triticum.
1 (bases 1 to 495)
Ogihara, Y. and Murai, K.
Expressed genes in Triticum aestivum
Unpublished (2002)
Contact: Tadao Shii-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshii@genes.nig.ac.jp.
Location/Qualifiers
1..495
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4365"
/clone="wh25m17"

FEATURES

source

/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4365"
/clone="wh25m17"

ORIGIN
/tissue_type="spike at heading date"
/dev_stage="Reekes" scale 10.5"
/clone_1lb="Y. Ogihara unpublished cDNA library, wh_h"

Query Match 74.4%; Score 18.6; DB 4; Length 495;
Best Local Similarity 84.0%; Pred. No. 7.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GAGGCAAGTAGTGCACACCTATGG 25
|||||
Db 269 GAGGCAAGTAGTGCACACATATGG 245
|||||

RESULT 14
BX499223/c 504 bp mRNA linear EST 04-SEP-2003
LOCUS BX499223 DKFZp779E0944.r1.779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION DKFZp779E0944_5', mRNA sequence.
ACCESSION BX499223 GI:32017467
VERSION BX499223.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 504)
Bioecker, H., Boecker, M., Mewes, H.W., Weil, B., Aml, C., Oeanger, A.,
Pobo, G., Han, M., and Wiemann, S.
EST (Bioecker, H., Boecker, M., Mewes, H.W., Weil, B., Aml, C., et al.)
Unpublished (2003)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysts, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No s1 sequence available.
This clone (DKFZp779E0944) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..504
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp779E0944"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="DH10B"
/clone_1lb="779 (synonym: hnccl)"
/note="Vector: pSPORT1_Sfi; Site_1: SfiI; Site_2: SfiIb"

FEATURES

source

ORIGIN
Query Match 74.4%; Score 18.6; DB 5; Length 504;
Best Local Similarity 84.0%; Pred. No. 7.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GAGGCAAGTAGTGCACACCTATGG 25
|||||
Db 496 GAGGCAAGTAGTGCACCTCCATATGG 472
|||||

RESULT 15
AM653346/c 506 bp mRNA linear EST 25-APR-2001
LOCUS AM653346 101999 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION AM653346
ACCESSION AM653346 GI:7419172
VERSION AM653346.1
KEYWORDS EST.

SOURCE
ORGANISM

Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.

REFERENCE
AUTHORS

1 (bases 1 to 506)
Smith, T.P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Cassas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C.,
Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,
Chicko-McKown, C.G., Pette, G., Holt, I., Karamycheva, S., Liang, F.,
Quackenbush, J. and Keel, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle

TITLE

JOURNAL
MEDLINE

Genome Res. 11 (4), 626-630 (2001)
2180013
11282978

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG

Plate: 104 row: P column: 14
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers

FEATURES

source

1..506
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH108"
/clone_id="WARC 180V"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

ORIGIN

Query Match 74.4%; Score 18.6; DB 2; Length 506;
Best Local Similarity 84.0%; Pred. No. 7.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAGGAGTAGTGGCCACACCTATGG 25
|||||
DB 86 GAGGAGCACTTGCACCTCCAGTGG 62
|||||

Search completed: December 15, 2004, 02:35:10
Job time : 81.7335 secs

This Page Blank (uspto)